LIGHT

A TEXTBOOK OF PHYSICS

By R. C. Brown, B.Sc., Ph.D.

Volume 1 MECHANICS AND PROPERTIES OF MATTER—contains the text of pp. 1–276

Volume 2 HEAT—contains the text of pp. 277–547

Volume 3 SOUND—contains the text of pp. 549–690

Volume 4 LIGHT—contains the text of pp. 693–1018

Volume 5 MAGNETISM AND ELECTRICITY *In preparation*

Complete Edition *In preparation*

LIGHT

by

R. C. BROWN, B.SC., PH.D.

Senior Lecturer in Physics
University College, London

With Diagrams

1724

LONGMANS, GREEN AND CO
LONDON · NEW YORK · TORONTO

LONGMANS, GREEN AND CO LTD
6 & 7 CLIFFORD STREET LONDON W I
BOSTON HOUSE STRAND STREET CAPE TOWN
531 LITTLE COLLINS STREET MELBOURNE

LONGMANS, GREEN AND CO INC
55 FIFTH AVENUE NEW YORK 3

LONGMANS, GREEN AND CO
20 CRANFIELD ROAD TORONTO 16

ORIENT LONGMANS LTD
CALCUTTA BOMBAY MADRAS
DELHI VIJAYAWADA DACCA

First published 1955

Printed in Great Britain by
NEILL & CO. LTD., EDINBURGH

PREFACE

Although in most printed syllabuses the section on Light occupies less space than some other branches of Physics, it is a fact that a comprehensive treatment of the subject tends to become quite lengthy. The fundamental principles of geometrical optics can be presented fairly briefly, but much of the value of an account of the subject lies in the discussion of various particular cases, in worked examples and, above all, in an abundance of ray diagrams. I have tried to make this volume useful in these respects as a source of reference for Intermediate and Advanced-Level students.

After much thought I have finally adopted the "real-is-positive" sign convention as being one of the most successful at this introductory level. It does, however, involve some supplementary rules (whereas the Cartesian convention does not), for example in connection with the formula for the power of a single refracting surface, and I have tried to make these as simple as possible. The use of $v/u = -m$, where m is negative when there is inversion, may seem at first sight to be unnecessarily complicated as compared with $v/u = m$, where m is negative when the object and image have different natures (real or virtual). I have used the former equation, however, on the grounds that if the latter is applied successively to two or more mirrors or lenses, then it is possible for the overall value of m, *i.e.* the product $m_1 m_2 \ldots$, to have, say, a negative sign when the initial object and final image are both real. This occurs, for example, in the case of two lenses in which the real image due to the first acts as a virtual object for the second and gives a real final image.

The theory of lenses has been developed from the properties of a single spherical refracting surface, as this method seems to me to be the most logical and rigorous method of approach and takes in its stride such cases as the thin lens separating two different media.

Some topics are included in this volume which, although they may not be specifically mentioned in the syllabuses, are nevertheless frequently touched upon in many teaching courses. They are indicated by footnotes in the text and include brief accounts of coma, curvature and distortion, the beginnings of thick-lens theory and qualitative treatments of various cases of diffraction.

My very sincere thanks are due to Mr. D. O. Wood for the many valuable suggestions he made when reading the volume in proof. The publishers' artist has again done good work in carefully copying my original drawings.

I wish to thank the following examining bodies for having given permission for the inclusion of questions from past papers:—

The Delegates of Local Examinations, Oxford (O.).
The University of Cambridge Local Examinations Syndicate (C.).
The University of London (L.).
The Joint Matriculation Board (J.M.B.).

The various examinations are indicated in the text by the above initials in conjunction with letters according to the following scheme:—

Intermediate—I.
Higher School Certificate—H.S.
Advanced Level—A.
Scholarship—Schol.
First Medical—Med.

In transcribing the questions I have usually refrained from modifying the various modes of writing the names of physical units. The reader is thereby enabled to become acquainted with other common methods of expressing units in addition to the "index" convention which is used in the text.

R. C. Brown.

CONTENTS

(The numbers refer to pages)

CHAPTER XLVIII

CHAPTER XLIX

CHAPTER L

CHAPTER LI

CHAPTER LII

CHAPTER LIII

CHAPTER LIV

CHAPTER LV

VOLUME IV

LIGHT

Chapter XLI

INTRODUCTION TO LIGHT

1. INTRODUCTION

The section of Physics which is known as Light concerns the nature and properties of the radiation which produces the sensation of sight. The student should remind himself at the outset of the relation between light and radiant heat which is discussed on page 527 of Vol. 2, and especially of the curves in Fig. 391 relating the energy and wave-length of the radiation given out by a black body at various temperatures. The wave-lengths which affect the eye lie between about 4×10^{-5} cm. (violet light) and $7{\cdot}2 \times 10^{-5}$ cm. (red). The radiation beyond the short wave-length end of this range is called **ultra-violet** and beyond the long wave-length end is the **infra-red.** It must be remembered that each of these is invisible, although ultra-violet radiation is often loosely referred to as "ultra-violet light."

It will be noticed that we are taking the wave nature of light for granted, and we shall continue to do so throughout most of this volume. The matter is not quite so simple as it might appear at first sight, however, and a short discussion of the nature of light is given at the end of the volume.

The speed of propagation of light *in vacuo* is the same as that of all other types of radiant energy, namely about 3×10^{10} cm. sec.$^{-1}$. The experimental determination of this constant is described in Chapter LV.

2. THE PROPAGATION OF LIGHT

Wave-Front.—If we imagine light waves to be given out in all directions by a point source, then a disturbance emitted at any given instant will, at some later time, have travelled a certain distance from the source in all directions. The surface on which lie all the points reached by a disturbance emitted from the source at a given instant is called a **wave-front.** In terms of oscillatory motion we may say that a wave-front is a surface joining points which are in the same phase. Evidently the wave-front due to a point source is spherical if the speed of propagation is the same

in all directions. If the medium is not isotropic, however, the wave-front will have different radii in different directions. This occurs in some crystals.

At a great distance from a point source in an isotropic medium the curvature of the spherical wave-front is so small that any small portion of it can be considered plane. We then have what is called a **plane wave** as distinct from a spherical wave.

Huygens' Principle.—This principle has already been mentioned in Vol. 3 (page 588) in connection with the propagation of sound. In effect, it is a geometrical construction by means of which the propagation of a wave-front may be represented.

In Fig. 490 the arc PQ represents part of a spherical wave-front which has emanated from a source S. According to Huygens' principle we regard every point on the wave-front as a source of **secondary wavelets.** If the wave-front PQ is spherical, then the medium is isotropic and the wavelets will also be spherical. Parts of a number of wavelets which were emitted from points on PQ at the same instant are shown in the drawing. The curve P′Q′ which is tangential to all the wavelets is the new position of the wave-front. A repetition of the same construction gives a later position of the wave-front and so on. If the wave-front is plane, PQ and P′Q′ are straight lines. Huygens' construction is of great use in discussing reflection and refraction, and especially the propagation of light in anisotropic crystals.

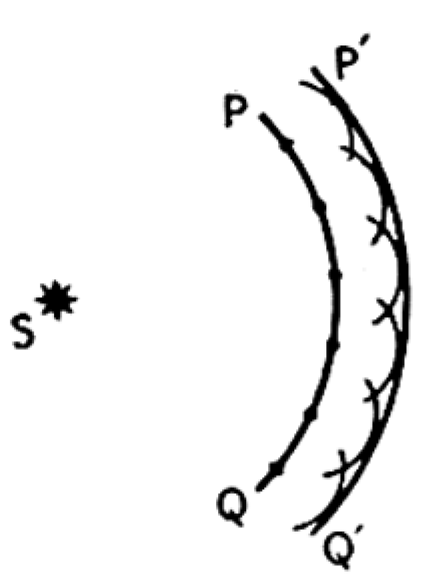


FIG. 490

Reflection.—We now apply Huygens' principle to the reflection of light. In Fig. 491 AB is a plane wave-front travelling towards a plane reflecting surface RR′. The parallel straight lines CA and DB which are perpendicular to AB represent the paths of the points A and B respectively and are **rays.** At the instant depicted, the point A has just arrived at the reflector. The reflected wave-front is obtained by considering the emission of wavelets back into the medium above RR′ from points on the wave-front as they arrive successively at RR′. Each wavelet represents the same phase, since all are given out from the same wave-front AB, and the reflected wave-front will therefore be the surface which is tangential to all the secondary wave-fronts. The first wavelet is given out from A and the last from B when this point arrives at E. When B has travelled to E the wavelet from A has travelled a distance AF equal to BE, and it is represented in Fig. 491 by the circular arc passing through F. For any point such as J between A and E the corresponding arc has a smaller radius JK

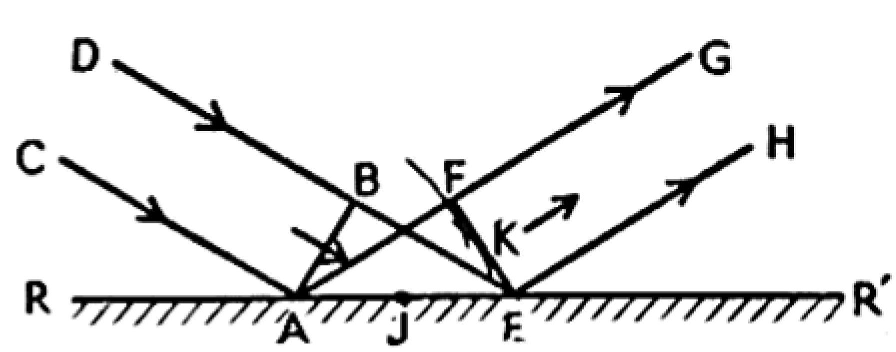


FIG. 491

because emission does not occur from J until the wave-front has reached that point. It is easy to see that the radius of a wavelet is proportional to the distance of its point of origin from E, and this means that the line EF which is tangential to the secondary wave-fronts, and therefore represents the reflected wave-front, is a straight line. The lines AFG and EH are perpendicular to EF and represent reflected rays.

The right-angled triangles BAE and FEA have the side AE in common and, as we have seen, BE = FA. They are therefore congruent and the angles $B\widehat{A}E$ and $F\widehat{E}A$ are equal. Thus the reflected wave-front makes the same angle with the reflecting surface as does the incident wave-front. This relationship is the well-known law of reflection. It is usually expressed by saying that corresponding reflected and incident rays make equal angles with the normal to the reflector at the point of incidence.

Refraction.—This is the name given to the change of direction which generally occurs when light passes from one medium to another in which its speed of propagation is different. In Fig. 492 AB represents a plane wave-front travelling in medium (1) and incident on the surface RR′ separating medium (1) from medium (2). The straight lines DBF and CA which are perpendicular to AB are incident rays. Suppose that a time t is necessary for the point B on the wave-front to arrive at F on the boundary. Thus $BF = c_1 t$, if c_1 is the speed of propagation in medium (1). During the same time t, the secondary wavelet which originated from the point A at the instant when the wave-front AB was in the position shown in the diagram has spread out into medium (2) until its radius AE is $c_2 t$, c_2 being the speed of the wave in medium (2). The wavelet from the other end of the wave-front (F) is now about to be initiated, while the wavelet from some point J between A and F has a radius intermediate between $c_2 t$ and zero. The wave-front in medium (2) is then represented by the line EF, which is tangential to all the wavelets and can easily be shown to be straight. The lines AG and FH are perpendicular to the wave-front and represent refracted rays.

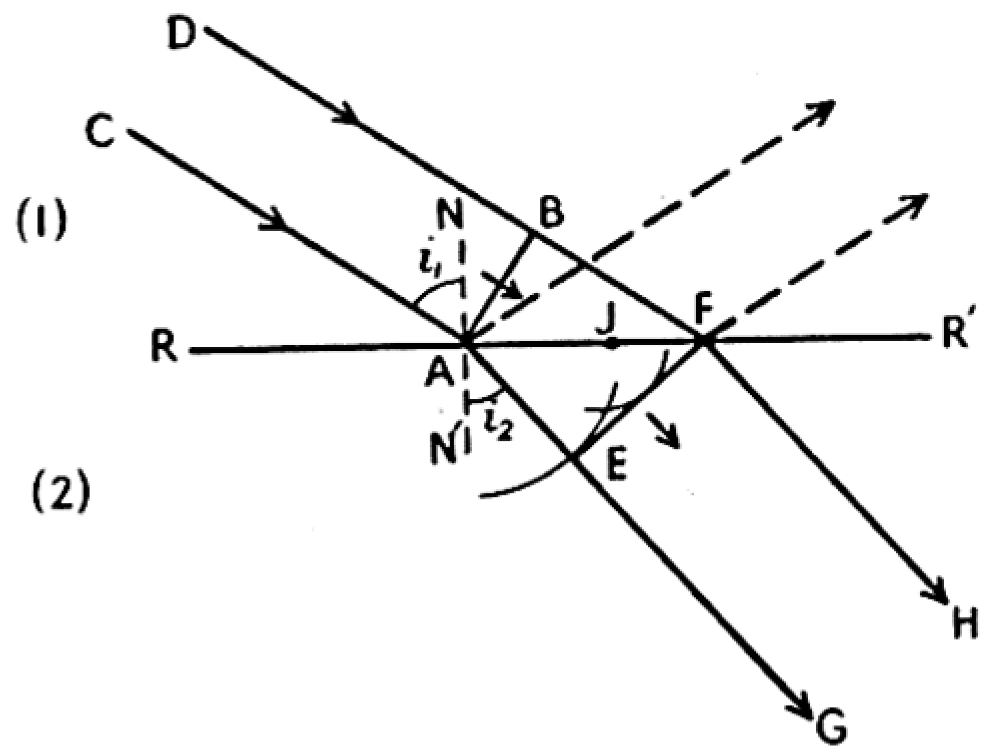


FIG. 492

The directions of the incident and refracted rays may be related by considering the angles i_1 and i_2 between the rays and the line NAN′ which is normal to RR′. These angles are called the angles of incidence and of refraction respectively. It is evident that i_1 is equal to the angle $B\widehat{A}F$ between the incident wave-front AB and the boundary RR′ since

AB is perpendicular to the rays. Similarly, i_2 is equal to $\widehat{AFE}$, the inclination of the refracted wave-front to the boundary. Since ABF and AEF are both right-angled triangles, we have

$$\sin i_1 = \sin \widehat{BAF} = \frac{BF}{AF}$$

and

$$\sin i_2 = \sin \widehat{AFE} = \frac{AE}{AF}$$

Therefore

$$\frac{\sin i_1}{\sin i_2} = \frac{BF}{AE} = \frac{c_1 t}{c_2 t} = \frac{c_1}{c_2}$$

It follows from this that the ratio of sin i_1 to sin i_2 is constant for a given pair of media, whatever the value of i_1. This relationship is known as **Snell's law** of refraction. The ratio of sin i_1 to sin i_2 is known as the **refractive index** for light travelling from medium (1) to medium (2). It will be noticed that in the example illustrated in Fig. 492 the fact that c_2 is less than c_1 has the consequence that i_2 is less than i_1, that is to say, the refraction causes the rays to be bent towards the normal. If c_2 is greater than c_1 the bending is away from the normal.

Rays of Light. Rectilinear Propagation.—It is often possible to describe accurately what we may call the "large scale" behaviour of light without taking into account its wave nature. We can imagine that a point source of light emits light energy equally in all directions, and that the energy travels away from the source along lines or rays which are straight if the medium surrounding the source is homogeneous and isotropic. The relationship between this idea of the propagation of radiant energy along straight lines and the conception of it as a wave motion is briefly discussed on pages 587–591 (Vol. 3) (in connection with sound), and a fuller discussion is to be found on pages 941–944, where it is shown that completely rectilinear propagation never really occurs. For example, the edges of the shadow of an obstacle are never absolutely sharp as required by the principle of rectilinear propagation, because the wave nature of light gives rise to **diffraction** at the edges of the obstacle. This causes a certain amount of light to enter the geometrical shadow, *i.e.* the region in which there would be complete darkness if the light energy travelled along straight lines. However, the small wave-length of light often renders the effects of diffraction so small that they can be disregarded, and the rectilinear propagation of light can then be adopted as a sufficiently accurate working hypothesis. The study of the behaviour of light with the help of this principle is known as **geometrical optics,** because it involves only the use of the principles of geometry combined with the geometrical relations which govern the reflection and refraction of light.

Point Source of Light.—It is often convenient in the study of light to use the idea of a point source of light, *i.e.* a source which has no size. This is, of course, a theoretical conception which cannot be realized in practice. A source of finite size (an **extended source**) can often be regarded as a collection of point sources in the same way as, in mechanics, a body can be treated as a collection of particles which have mass but no size.

Pencils of Light.—It is not possible in practice to produce a single ray of light. As has been already pointed out, rays are a theoretical conception which enables us to represent the behaviour of light. When light from a source O (Fig. 493 (i)) passes through a hole in a screen, the

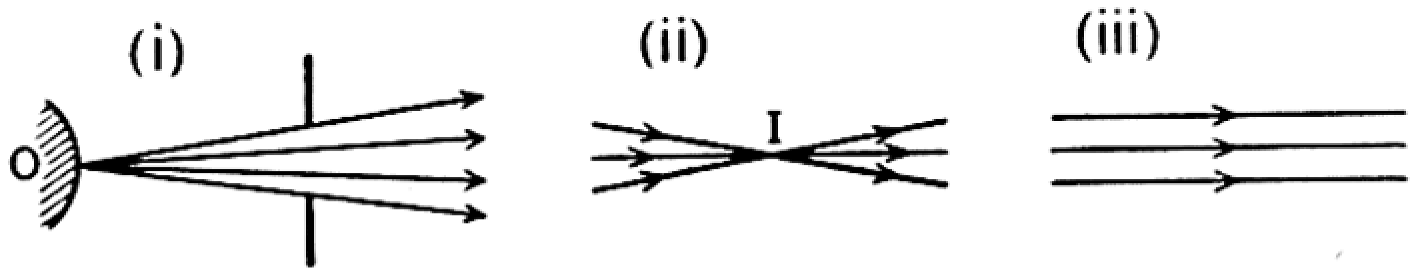


FIG. 493

passage of the luminous energy from any particular point on O can be represented by rays as shown in the figure. Such a collection of rays is called a **diverging pencil.** There are, of course, an infinite number of rays in the pencil although only four are shown. It is often only necessary to draw the extreme rays. In practice all the other points on the object cause pencils of rays to issue from the hole in different directions, and it is impossible to obtain a single pencil with a real source of light. A **converging** pencil of light is shown in Fig. 493 (ii). It will be noticed that after the converging rays have crossed at the point I (which is called their **focus**) they become a diverging pencil.

A **parallel** pencil of light is shown in Fig. 493 (iii). Converging and parallel pencils consisting of rays originating from a single point cannot be formed without the action of mirrors or lenses. Rays originating from a point source are always diverging. Since pencils of light rays require a point source in the first place they cannot be realized in practice. Nevertheless, we shall make frequent use of them when we treat real sources as collections of point sources.

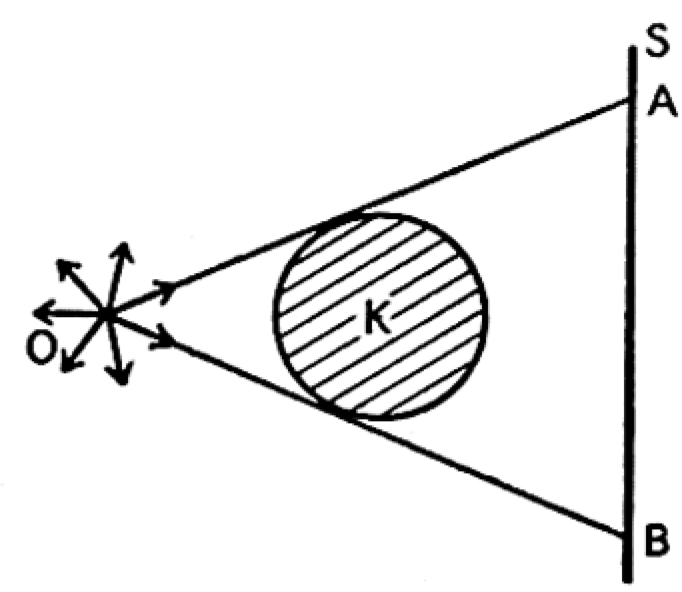


FIG. 494

Shadows.—A very simple example of the application of geometrical optics is afforded by the formation of shadows. The only principle involved is the rectilinear propagation of light.

In Fig. 494, let O be a point source of light and let K be an opaque obstacle which, for the sake of simplicity, we suppose to be a sphere. Light energy flows in all directions from O along straight-line rays as indicated

in the figure (*cf.* the lines of electric or magnetic force from a point charge or pole). The rays are unable to pass through K so that a shadow is formed on the screen S. The outline of the shadow (AB) is determined by drawing rays from O which touch K and continue to S. The shadow would then be a region of uniform darkness (supposing that light from O is not reflected on to the screen by any other near objects), and it would be separated by a sharp line from the surrounding region of the screen in which the illumination is exactly the same as if K were absent. Although diffraction of the light waves at the boundary of the obstacle does in fact cause a slight blurring of the edges of the shadow, this is scarcely noticeable in ordinary circumstances, and we can say that sharp uniform shadows similar to this are formed in practice by very small sources of light, especially when the obstacle is at a considerable distance from the source, the finite size of which is then relatively less important. But in fact no real light source can be said to be a point source, and we next consider the type of shadow which is formed by a uniform extended source. This is shown in Fig. 495, where the source is O_1O_2, the obstacle K and the screen S. Four straight lines are drawn from O_1 and O_2 to touch K in the way shown and are then continued to the screen S. Evidently no light whatever (apart from diffraction effects) can reach the screen in the region between A_2 and B_2, which is therefore a complete shadow and is known as the **umbra.** As we go outwards from this region the illumination of the screen increases. For example, the point y can receive light from points on the source between O_1 and x. Thus the umbra is surrounded by a region of partial shadow known as the **penumbra,** in which the illumination increases progressively until at the edges A_1 and B_1 it becomes the same as it would be if K were absent.

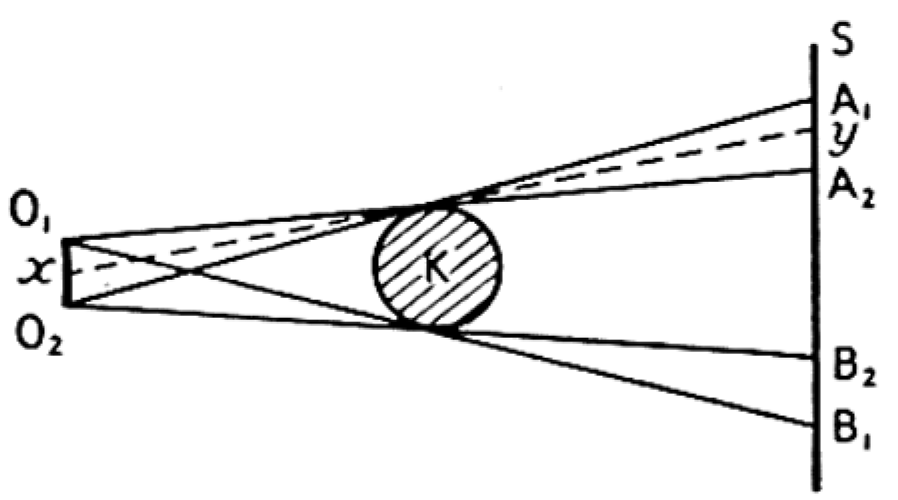


FIG. 495

The Pinhole Camera.—This device (Fig. 496) also illustrates the rectilinear propagation of light. It consists of a box with a small hole P in one side, the opposite side S being covered with a photographic film or plate. Alternatively S may be made of ground glass so that the performance of the camera may be observed while it is in action. Let O_1O_2 be an object which is either self-luminous or illuminated by a separate source of light. Of the light which leaves the bottom of the object (O_2) in all directions only those rays within the narrow cone indicated in the figure will enter the hole P.

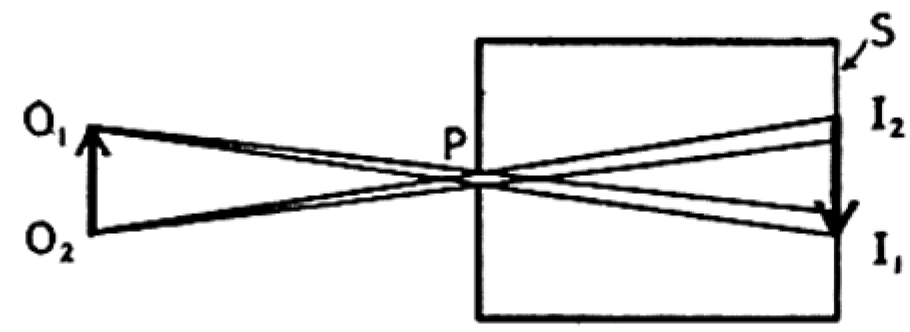


FIG. 496

These will produce a luminous patch I_2 on S of the same shape as the hole. The smaller the hole the smaller is the patch (although diffraction effects begin to be noticeable when a very small aperture is used). Every point on the object forms a similar area of illumination on S, with the result that a picture or **image** of O_1O_2 is formed on S between the points I_1 and I_2. The image can never be perfectly defined, because each point of the object gives rise to a finite area on the image and consequently there is overlapping and blurring. The image is made sharper but less bright by reducing the size of the hole, provided the stage at which diffraction occurs is not reached. There is no necessity for focusing the camera as there is when a lens is used. It should be noticed that the image is inverted, and that the ratio of its transverse dimensions to the corresponding dimensions of the object is equal to the ratio of the distance of the image from the hole to that of the object from the hole.

EXAMPLES XLI

1. Explain Huygens' principle, and show how it can be used to establish the laws of reflection and refraction of light.

2. Describe simple experiments to demonstrate the rectilinear propagation of light in a homogeneous medium.
Give a diagram to explain the action of a pinhole camera, and discuss the effect of increasing the size of the hole.
To what do you attribute the small circular patches of light sometimes observed in clear weather on the ground under trees? (L.Med.)

Chapter XLII

REFLECTION BY PLANE SURFACES

1. THE LAWS OF REFLECTION

A Light Ray Striking a Boundary.—In Fig. 497 AB represents a light ray travelling in medium (1) and striking the surface of a different medium (2). Some of the light energy of the **incident ray** (*i.e.* AB) is always reflected by the surface, and in the figure this reflected ray is indicated by BC. The remaining light energy passes into medium (2), usually undergoing a change of direction which is known as **refraction.** The refracted ray is denoted by BD. Its intensity diminishes on account of absorption as it travels through the medium. With opaque substances the refracted light reaches zero intensity after a very short (but finite) distance. Very thin sheets of materials such as metals, which we usually regard as opaque, will transmit a certain amount of light.

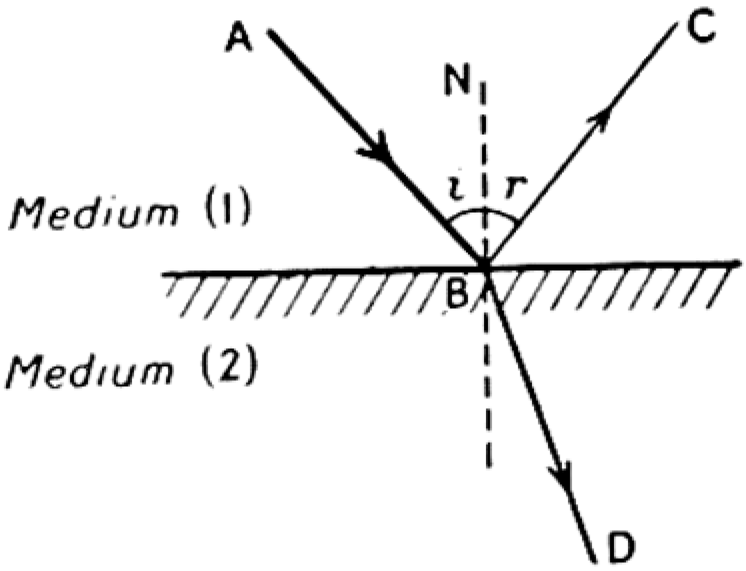


FIG. 497

In dealing with reflection we shall frequently refer to the reflecting surface as "the mirror," not necessarily implying, however, that it is of the familiar type. An unsilvered glass surface can act as a mirror, although it is not very efficient.

The Laws of Reflection.—The process of reflection is considered from the point of view of the wave nature of light on page 694, where the second law stated below is deduced. We now state the laws as experimental facts.

The line BN in Fig. 497 is drawn perpendicular to the direction of the mirror at B and is known as the **normal** at the point of incidence. The angle which the incident ray makes with the normal is the **angle of incidence** (i), and the angle between the normal and the reflected ray is the **angle of reflection** (r).

The two laws of reflection state that

(i) The incident ray, the reflected ray and the normal to the mirror at the point of incidence all lie in the same plane.

(ii) The angles of incidence and reflection are equal.

Experimental Verification.—Many of the principles of geometrical optics can be investigated by marking the directions of light rays with

pins. In order to test the laws of reflection, a reflecting surface such as the polished surface of a block of glass is required. (A glass mirror silvered on the back introduces a complication on account of refraction at the front surface—see page 751.) The glass block is placed on a horizontal piece of drawing-paper (Fig. 498) so that the surface which is to be used as a mirror is vertical. The lower edge of the surface is marked by a line M_1M_2 drawn on the paper. Two upright pins P_1 and P_2 are then stuck in the paper about 6 inches apart so that the line joining them makes a convenient angle with the reflecting surface. On looking into this surface, P_1 and P_2 will be seen by reflection as the images P_1' and P_2'. Using one eye only, the head is moved until P_2' covers P_1', and a third pin P_3 is placed so that it is in line with both the images. Finally, a fourth pin P_4 is placed about 6 inches from P_3 to be in line with P_1', P_2' and P_3.

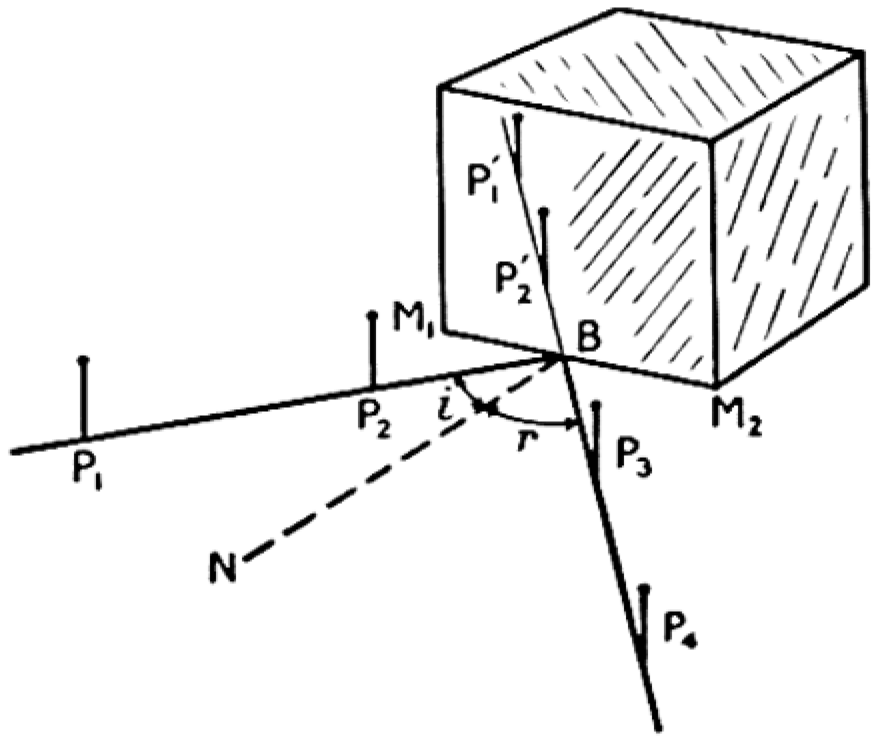


FIG. 498

The glass block is then removed, the pins are taken out of the paper, and the holes which they leave are ringed with small circles and marked P_1, P_2, etc. The points P_1 and P_2 are joined by a straight line and so are P_3 and P_4. These two lines, when produced, should meet at B on the line M_1M_2. The normal BN is drawn at this point, and the angles i and r should, when measured, be found to be equal to each other to within, say, half a degree. These are evidently angles of incidence and reflection, because a ray of light travelling from P_1 to P_2 and thence to the reflecting surface is reflected in the direction P_3P_4. Thus the second law is verified. The first law can be seen to be obeyed because if the pins are all of the same height, the heads of P_3 and P_4 are observed to be in line with those of the images P_1' and P_2'.

Effect of Rotation of the Reflecting Surface.—Let a ray of light AB (Fig. 499) be incident on a reflecting surface or mirror M_1M_2 with an angle of incidence i. The reflected ray is BC and the angle of reflection is r. Then let the mirror be moved to a new position $M_1'M_2'$, which makes an angle θ with its original position. If the direction of the incident ray AB remains unaltered, it will now strike the mirror at B′ and be reflected in the direction B′C′. If the new angle of incidence is i' and the angle of reflection r', we have $i=r$ and $i'=r'$. The new normal B′N′ is inclined at an angle θ to the direction of the original normal BN, because this is the angle through which the mirror has turned. Therefore i' is greater than i by an angle θ. Thus

$$i'=i+\theta \quad . \quad . \quad . \quad . \quad . \quad (1)$$

But BC makes an angle of $(i+r)$ or $2i$ with AB, while B′C′ makes an angle of $(i'+r')$ or $2i'$ with AB. Therefore the reflected ray has been turned through an angle of $(2i'-2i)$, which, by equation (1), is equal to 2θ. Thus when the direction of the incident ray remains constant and the mirror is rotated, the reflected ray is caused to rotate through twice the angle of the mirror. This principle is involved in the working of the sextant, the optical lever (Vol. 2, page 318), and in all cases where the rotation of a body is measured by the reflection of light from a mirror attached to the body, *e.g.* the mirror galvanometer.

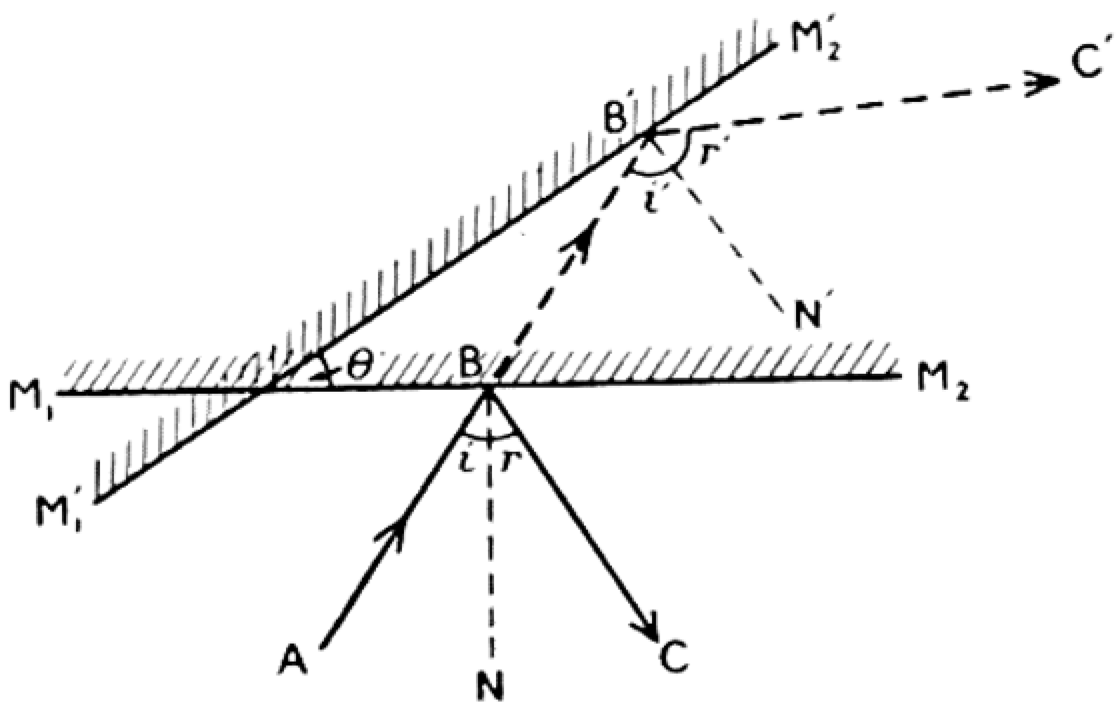


FIG. 499

***Example.**—Derive a formula for the total deviation of a ray of light when it is reflected by two plane mirrors which are inclined to each other at an angle θ, the plane of incidence being at right angles to the line of intersection of the mirrors.*

In Fig. 500, AM_1 and AM_2 are the two mirrors whose planes intersect at an angle θ at A. The ray is first incident on AM_1 at the point B, the angle of incidence between the ray and the normal BD being i_1. The angle of reflection is also i_1 as marked on the diagram. Thus the deviation from its original direction which the ray suffers at B is $(180^\circ - 2i_1)$. Similarly at C on the second mirror, where the angle of incidence is i_2, the deviation is $(180^\circ - 2i_2)$. The total deviation is therefore $(360^\circ - 2i_1 - 2i_2)$ or $\{360^\circ - 2(i_1+i_2)\}$.

The angle between the normals is θ, as marked in the diagram, so that, considering triangle CBD, we have $\theta = i_1 + i_2$.

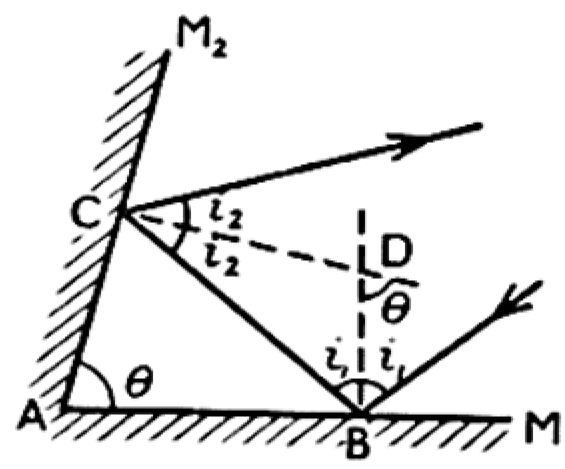


FIG. 500

Therefore the total deviation is equal to

$$360^\circ - 2\theta$$

It should be noticed that this deviation is independent of the angle of incidence.

The Sextant.—This instrument depends on plane reflection, and is used for measuring the angle subtended at the observer by two distant objects, and especially for measuring the elevation of the sun above the horizon when taking bearings at sea. The instrument is shown diagrammatically in Fig. 501 (i). It contains two plane mirrors. The mirror M_1 is rigidly fixed to the frame of the instrument and only half of its surface is silvered. M_2 can be rotated about an axis which is shown as a dotted line in the drawing, and its angle of rotation can be measured by a pointer attached to the mirror mounting and moving over a circular scale fixed to

the frame of the instrument. The observer places his eye at a small hole or telescope and receives light direct through the non-silvered part of M_1 in one half of the field of view, and by reflection in M_2 and M_1 in the other half of the field. It is possible, by rotating M_2, to see the same object in the two halves of the field and to make the two images continuous with each other. The angle subtended at the instrument is zero in this case, and the pointer reading is the zero of the sextant. This setting is shown in Fig. 501 (ii). To find the angle subtended by any two distant

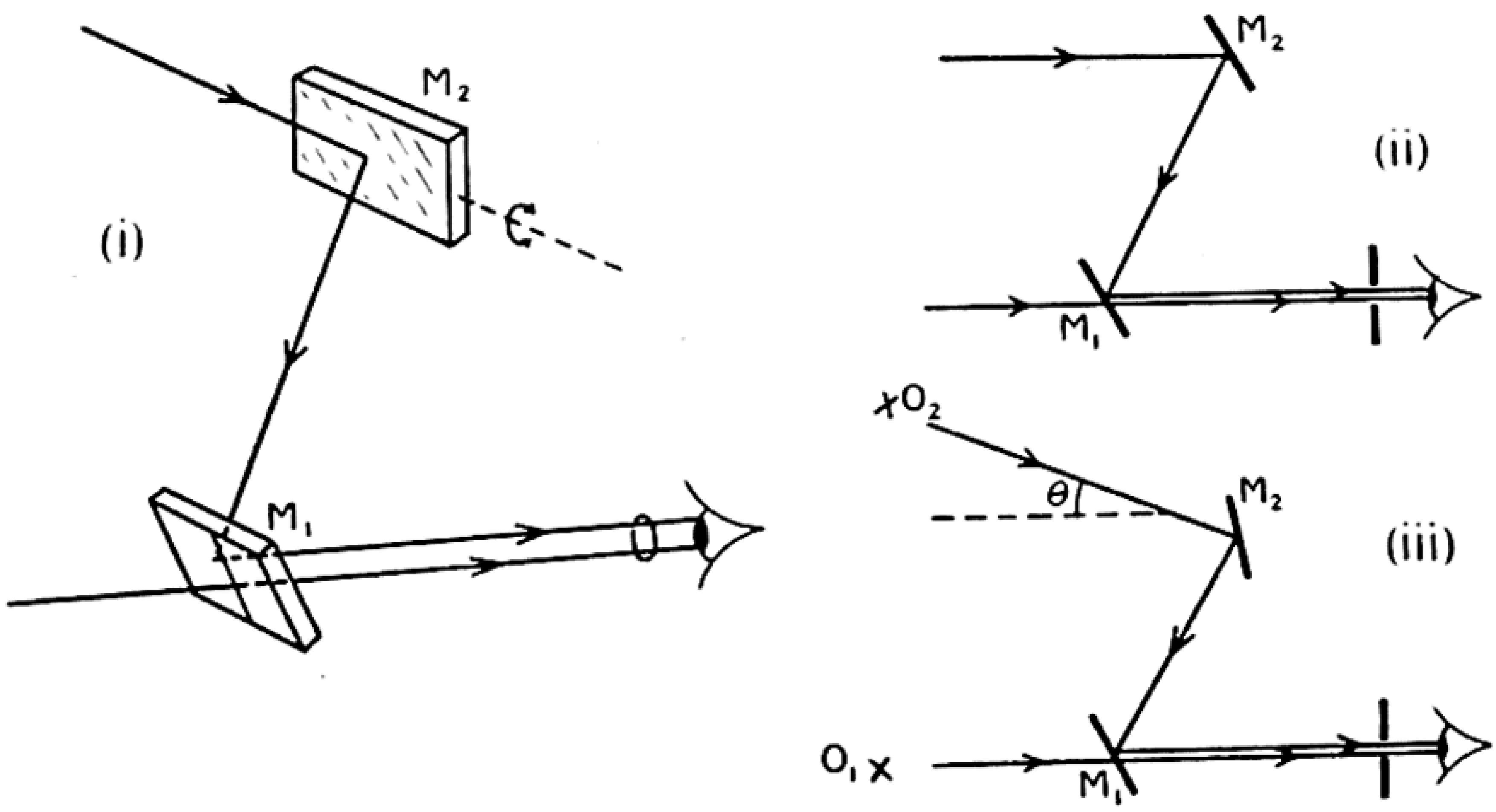


FIG. 501

objects, the observer directs the instrument so as to see one of them (O_1) through the unsilvered glass and rotates M_2 until the second object (O_2) is seen in the other half of the field alongside the first (Fig. 501 (iii)). If θ is the angle subtended by O_1 and O_2 at the instrument, it is evident that M_2 had to be rotated through $\theta/2$ from its zero position in order to give the reflected ray the same direction as it had when the zero setting was made. The difference between the new pointer reading and the zero reading is therefore equal to half the angle θ. In many instruments the circular scale is graduated in degrees which are half their proper size so as to make an automatic multiplication by 2.

2. THE FORMATION OF IMAGES BY PLANE MIRRORS

Image of a Point Object Due to a Single Mirror.—When objects are "seen in a mirror," the eye is really receiving reflected rays whose directions are what they would be if the mirror were absent and the real objects themselves were situated where they appear to be when viewed in the mirror. The mirror is said to form an image of the object, and we now proceed to examine this phenomenon from the point of view of the laws of reflection.

In Fig. 502, a point object O is placed so that rays of light from it such as OM_1 and OM_2 strike a plane reflecting surface on which M_1 and M_2 are situated and give rise to reflected rays M_1P_1 and M_2P_2. These reflected rays are produced backwards to meet at I, and the straight line IO crosses the reflecting surface (produced if necessary) at Q.

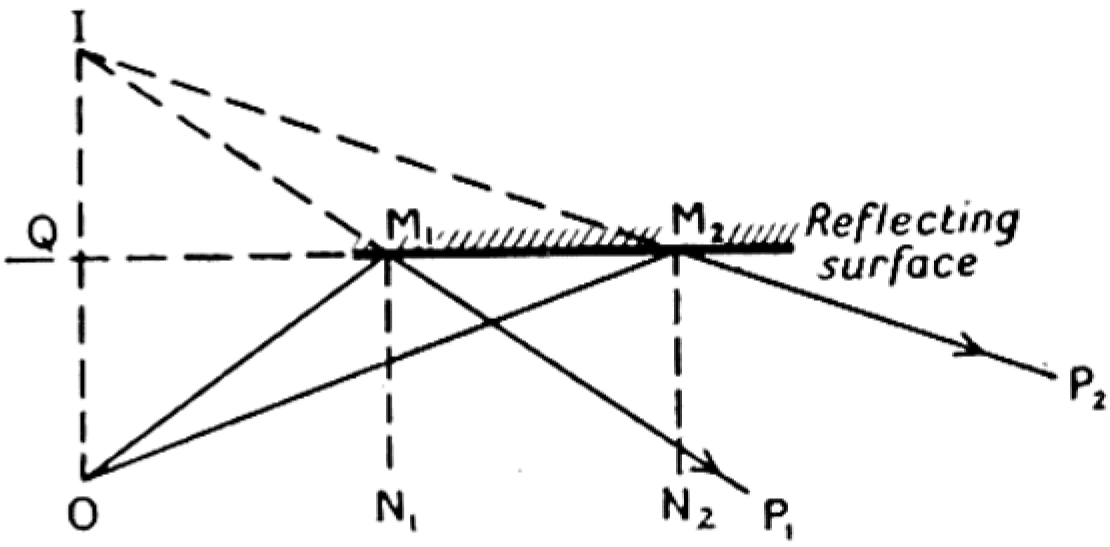


FIG. 502

The lines M_1N_1 and M_2N_2 are normals, and it is quite easy to use the laws of reflection and simple geometry to show that $I\widehat{M_1}M_2 = O\widehat{M_1}M_2$ and $I\widehat{M_2}M_1 = O\widehat{M_2}M_1$. Therefore the triangles OM_2M_1 and IM_2M_1 have a common side (M_1M_2) and the adjacent angles are equal, so that they are congruent, and

$$M_1I = M_1O$$

Thus the triangles IM_1Q and OM_1Q have the above two sides equal, QM_1 is common to both triangles, and the angles $I\widehat{M_1}Q$ and $O\widehat{M_1}Q$ are equal. The triangles are therefore congruent, so that

$$IQ = OQ$$

and

$$M_1\widehat{Q}I = M_1\widehat{Q}O = 90°$$

Therefore the point I is on the straight line passing through O and perpendicular to the reflecting surface, and is as far behind the surface as O is in front of it. This complete statement of the position of I has been arrived at by considering *any* two pairs of incident and reflected rays, and it is therefore true of *all* rays from O which strike the mirror. All reflected rays, therefore, diverge from the same point I. An eye which receives these reflected rays sees an image of the point O situated at I. Fig. 503 illustrates this by showing the reflection of two divergent pencils from O and their reception by an eye placed successively in two different positions. The position of the image is independent of that of the eye. In order to construct such a diagram the position of I is first marked

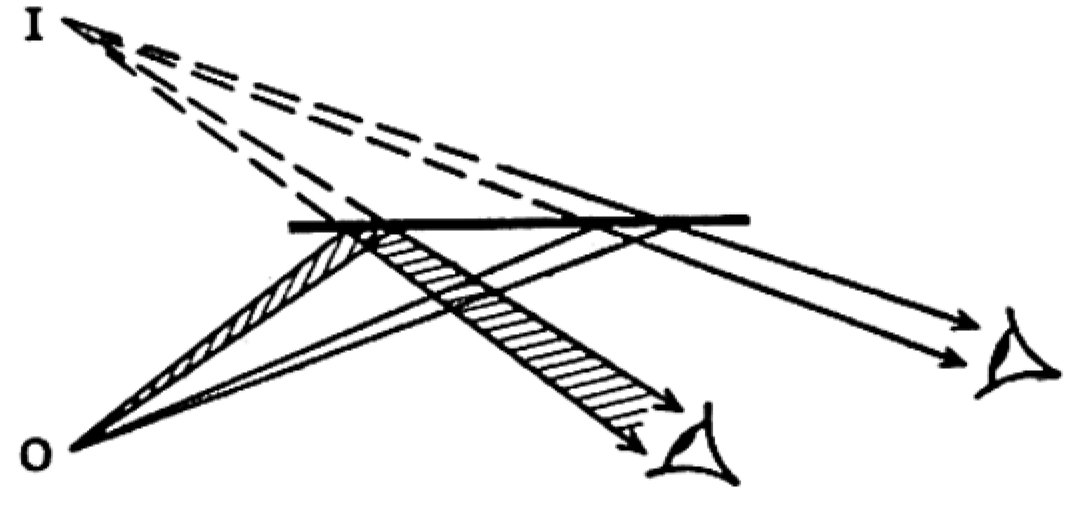


FIG. 503

on the normal as far behind the mirror as O is in front. Then any incident ray is drawn from O to the mirror. A dotted line is next drawn from I to the point of incidence and produced in front of the mirror as a full line to represent the reflected ray. Any pair of such rays marks the boundaries of a pencil of rays diverging from O.

Image of a Finite Object.—In Fig. 504, O_1O_2 is the outline of a surface which is placed in front of a mirror. Its image is I_1I_2, each point on the object forming a point on the image (*e.g.* O and I) which lies on the perpendicular as far behind the mirror as the point on the

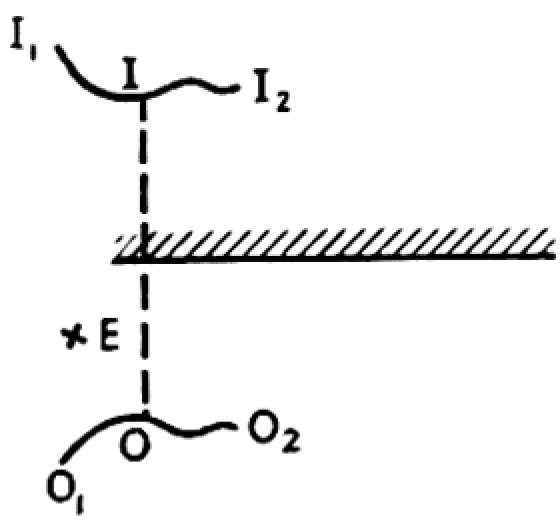


FIG. 504

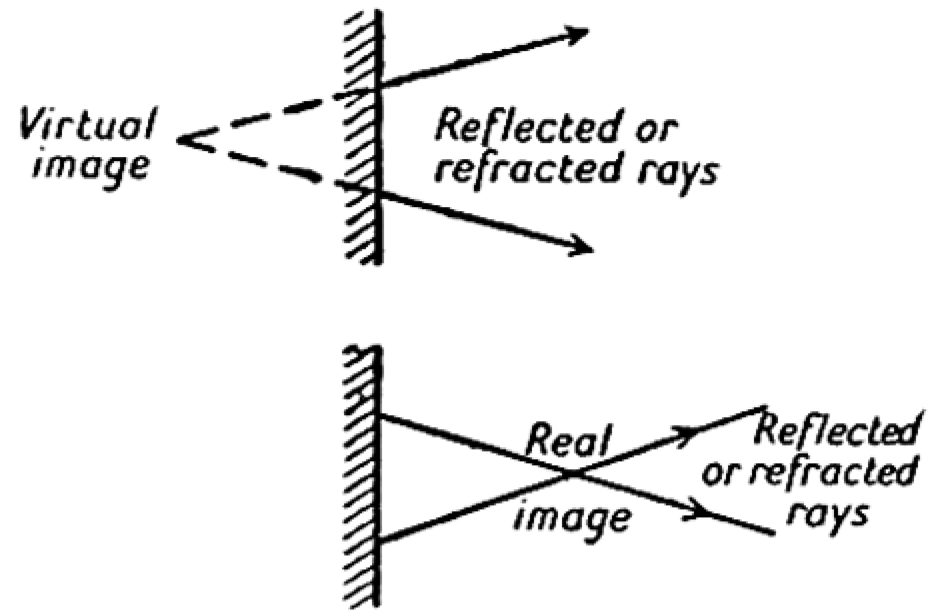


FIG. 505

object is in front of it. From this we see immediately that the image is the same size as the object.

Although the image is **upright** or **erect** (*i.e.* not inverted, as are some images which we shall discuss later on) it is said to be **laterally inverted,** *i.e.* it is reversed from left to right. If an eye is placed at E and views the object direct it sees O_1 on the right-hand side of the object; whereas if the eye is turned to observe the image in the mirror then I_1, which is the image of O_1, is on the left-hand side of the image. This is the type of reversal which occurs on blotting-paper when it is used to dry writing—hence the method of reading the marks on blotting-paper by holding it in front of a mirror.

Virtual and Real Images.—It is very important to understand what is meant by the statement that the image formed by a plane mirror is **virtual.** It signifies that the reflected rays by means of which the eye sees the image do not actually proceed from the image. This means that no light from the object goes to the image. The other type of image which we shall encounter later on is a **real** image. In this case the light rays, after being reflected or refracted, actually converge to the image and, having crossed over, diverge from it. The distinction between the two types of image is illustrated in Fig. 505, where the vertical straight line represents an image-forming device such as a mirror or lens. The original incident rays coming from the object are not shown. It will be realized that a virtual image is due to rays which diverge on leaving the mirror or lens, while a real image is produced by rays which converge.

A real image can be formed on a screen (*e.g.* in a cinema) while a virtual image cannot.

An ordinary object can never give rise to a real image in a plane mirror, but if light from the object is first made to converge (*e.g.* by means of a converging lens) before striking the mirror a real image will be produced (Fig. 506). In this case we can regard O, the point to which the incident light is converging, as the object; and, because the light does not actually originate from O, it is called a **virtual object.**

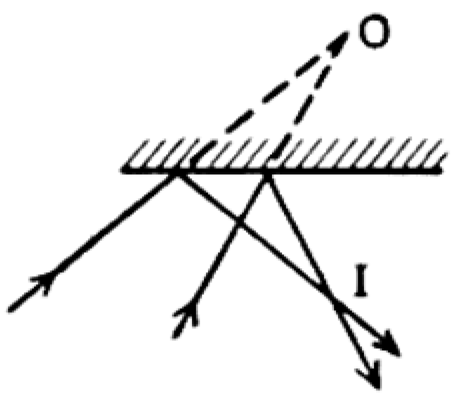


FIG. 506

Summary.—The complete specification of the image of a real object formed by a plane mirror is as follows:—

Nature.—Virtual, erect, laterally inverted.

Position.—On the perpendicular from the object to the mirror, as far behind the mirror as the object is in front of it.

Size.—Same as object.

3. IMAGES FORMED BY MULTIPLE REFLECTIONS

Inclined Mirrors.—Consider a point object, O (Fig. 507), situated in front of two reflecting surfaces. If this arrangement is set up, a number of images of O can be observed, and we proceed to investigate their formation as follows. Let the mirrors be referred to by the numbers (1) and (2) as shown in the drawing. The images I_1 and I_2 are formed by

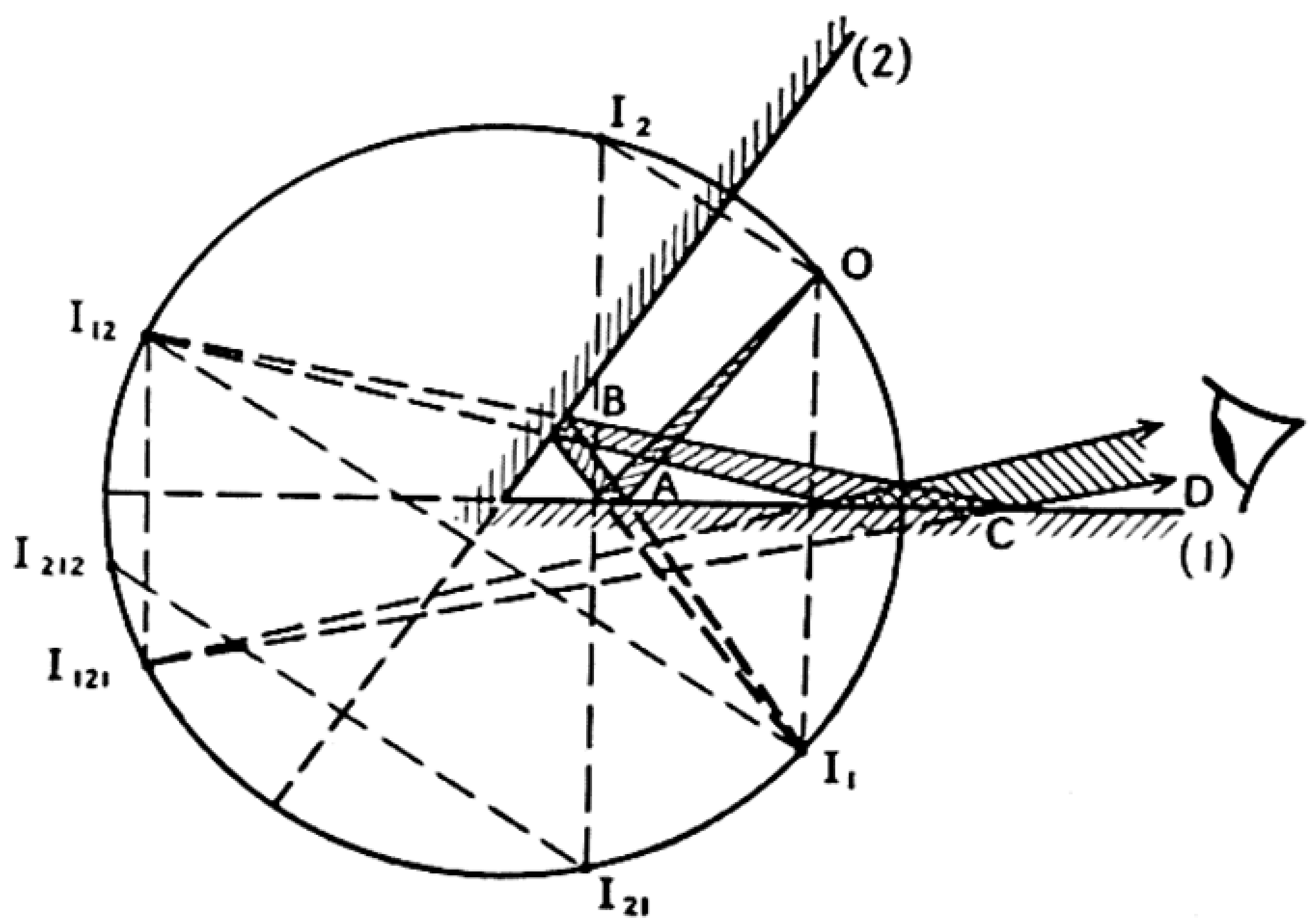


FIG. 507

mirrors (1) and (2) in the ordinary way. The light rays reflected from mirror (1) which give rise to the image I_1 are diverging from I_1, and when they strike mirror (2) they will be reflected by it so as to produce another image, just as they would if I_1 were an actual object. This image is I_{12}

and its distance behind mirror (2) is equal to the distance of I_1 in front of it. Similarly, I_2 acts as an object for mirror (1) and produces the image I_{21}. By the same principle both these secondary images give rise to further images, and the process of image formation finishes only when an image falls behind both the reflecting surfaces. The dotted perpendiculars are all chords of a circle which passes through O and has its centre at the point of intersection of the two lines representing the mirrors. In constructing a diagram such as that in Fig. 507, it is very helpful to draw this circle first because all the images lie on its circumference.

Having found the positions of the images as described above, we can now draw the path of the rays by which a suitably placed eye would see any particular image. Suppose we choose I_{121}. The last reflection by which this image was formed took place in mirror (1), so that the eye must look into this mirror in order to see it. A dotted line is therefore drawn from I_{121} in the direction of the eye, and continued as a full line to represent a light ray in front of the mirror at C. The ray CD is produced by the reflection of light coming from the direction of I_{12}, so that a line is next drawn from I_{12} to C crossing mirror (2) at B. Then I_1 is joined to B by a line which crosses mirror (1) at A, and finally A is joined to O. In order to construct a pencil of rays another ray entering the eye is traced back to O in the same way.

With certain simple angles of inclination of the mirrors (*e.g.* 90° and 60°), the last two images which fall behind both reflecting surfaces coincide with each other.

Parallel Mirrors.—When two mirrors are arranged with their reflecting surfaces facing and parallel to each other, an object placed between them forms an infinite number of images although their intensity diminishes with the successive reflections. The principle is exactly the same as in the previous case. The placing of the images on a drawing is facilitated by constructing a series of lines as shown in Fig. 508 parallel to the

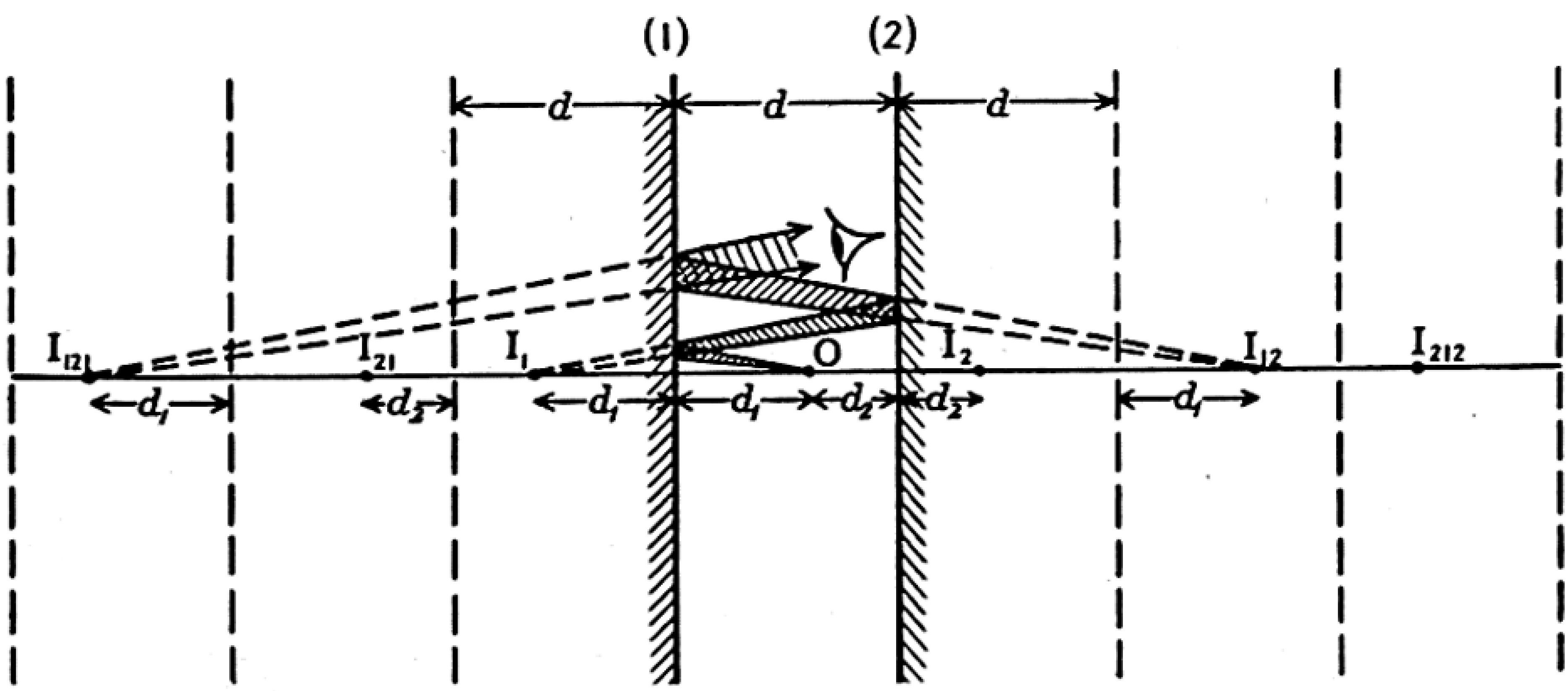


FIG. 508

mirrors, adjacent lines being separated from each other by a distance equal to the separation of the mirrors (d). If the object O is at a distance of d_1 from mirror (1) and d_2 from mirror (2), the first image of O due to mirror (1) will be I_1, at a distance d_1 behind the mirror (1). This image acts as an object for mirror (2) and produces I_{12}, at a distance $(d_1 + d)$ behind it. The next image in this series is I_{121}, at a distance $(d_1 + 2d)$ behind mirror (1), and so on. The other series of images begins with I_2, which is d_2 behind mirror (2), and continues with I_{21}, I_{212}, etc. The pencil of rays shows how an eye looking into mirror (1) sees the image I_{121} by successive reflections. All the images of a finite object are of the same size as the object itself, although they appear to an observer situated between the mirrors to become smaller and smaller as they recede from the eye—just as the size of the original object would diminish if it were placed successively in the positions of the images.

EXAMPLES XLII

1. (*a*) State the laws of reflection of light.
   (*b*) Find an expression for the deviation produced when a ray of light, incident at an angle of i, is reflected by a plane mirror.
   (*c*) Show that when such a mirror is turned through an angle θ about an axis in its plane, a ray of light, incident at *any* point on the mirror, is turned through an angle 2θ.
   (*d*) Of what practical importance is the result expressed in (*c*)? (L.Med.)

2. Two plane mirrors are inclined to each other at a fixed angle. If a ray travelling in a plane perpendicular to both mirrors is reflected first from one and then from the other, show that the angle through which it is deflected does not depend on the angle at which it strikes the first mirror.

   Describe and explain the action of either a sextant or a rear reflector on a bicycle. (L.H.S.)

Chapter XLIII

REFLECTION BY SPHERICAL MIRRORS

1. IMAGES OF POINT OBJECTS

Definitions.—A spherical mirror is a reflector whose reflecting surface is part of the surface of a sphere. The mirror is said to be **concave** or **convex** according as the reflecting surface is concave or convex towards the space in which the incident light is travelling. The centre of the sphere of which the mirror is a part is the **centre of curvature** of the mirror, and we shall denote this point by the letter C (Fig. 509). The radius of the sphere is the **radius of curvature** of the mirror.

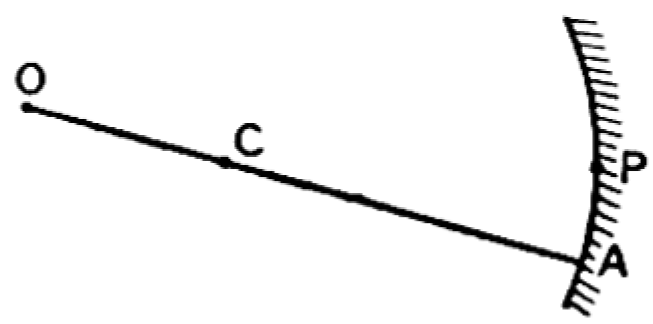


FIG. 509

In Fig. 509, in which the mirror is concave, a point object O is situated in front of the mirror, and the straight line passing through O and C is called the **principal axis** *for this particular point object.* The point A, where the principal axis cuts the reflecting surface, is the **pole** corresponding to the object O.

It is evident from the above definitions that there is no such thing as *the* principal axis or *the* pole of a spherical mirror. The locations of these are determined by the position of the particular point object we are considering. It is common practice in elementary optics, however, to regard the central point of the reflecting surface as *the* pole of the mirror (P in Fig. 509), although this point has no unique position as regards the actual spherical surface—only as regards the edges of the mirror itself. The line through the pole P and the centre of curvature C is then called *the* principal axis. We shall adopt this practice.

Concave Mirror.—*Case* 1. In Fig. 510, O is a point object situated on the principal axis of a concave mirror. Consider any incident ray such as OA. The normal at A is the radius AC, and the reflected ray AI is in such a direction as to make the angles of incidence and reflection equal. Each of these angles is denoted by i on the diagram. The point I is, as yet, simply the point at which this particular reflected ray crosses the principal axis PC, and we shall show that, subject to a very important condition, *all* the rays from O which strike the mirror are reflected so as to pass through I.

Let the angles indicated in the diagram have magnitudes α, θ and β in circular measure. Then, since the exterior angle of a triangle is equal to the sum of the two opposite interior angles, we have

$$\beta = i + \theta \text{ in triangle AIC}$$

and

$$\theta = i + \alpha \text{ in triangle ACO}$$

Elimination of i from these two equations gives

$$\beta + \alpha = 2\theta \quad . \quad . \quad . \quad . \quad . \quad (1)$$

In order to test whether the position of I depends on the choice of incident ray from O, it is necessary to find an expression for the distance of I from, say, P. This is a complicated matter unless we confine ourselves to cases in which the distance AP is small compared with the distance of I, C and O from P. This means that the angles α, β and θ are small, and that the rays are almost parallel with the principal axis. Such rays are said to be **paraxial.**

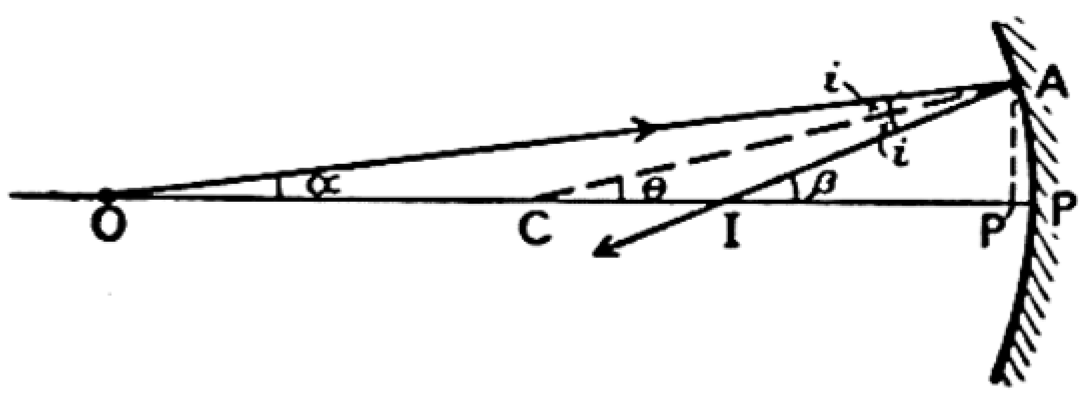


FIG. 510

Let AP′ be perpendicular to the principal axis. Then

$$\sin \alpha = \frac{AP'}{AO}$$

If α is small its sine is equal to its magnitude in radians, so that

$$\alpha = \frac{AP'}{AO} \quad \text{approximately}$$

Also

$$AO = OP \quad \text{approximately}$$

so that

$$\alpha = \frac{AP'}{OP} \quad \text{approximately}$$

Similarly

$$\beta = \frac{AP'}{IP} \quad \text{approximately}$$

and

$$\theta = \frac{AP'}{CP} \quad \text{approximately}$$

Thus, for paraxial rays, equation (1) becomes

$$\frac{AP'}{IP}+\frac{AP'}{OP}=2\times\frac{AP'}{CP}$$

or

$$\frac{1}{IP}+\frac{1}{OP}=\frac{2}{CP}$$

It is important to realize the significance of this result apart from the formula itself. The length AP′ has been cancelled and does not appear in the equation, which means that the equation is true for all positions of A provided that AP′ is small enough to allow the above approximations to be made. Thus, for given values of OP and CP, the value of IP is independent of the position of A; which means that all paraxial incident rays from O are reflected so as to pass through I, and I is the **image** of O. A number of incident and reflected rays are shown in Fig. 511. Of course, the paraxial condition is never perfectly obeyed, but in practice we find that if the width of the mirror is, say, 2 cm. and the distances of C and O from P are not less than 10 cm., then a clear image of O is formed at I. What happens when the rays are not paraxial will be mentioned later.

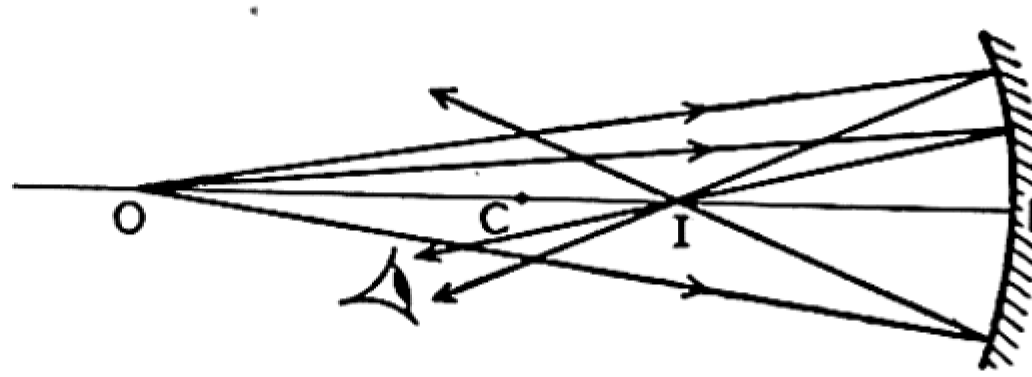


FIG. 511

It will be noticed that the image is formed on the same axis as the object. Light incident from O in the direction of the axis is, of course, reflected back along its own path.

The image is a **real** image because light from the mirror is converging towards it and actually passes through it. If O were a point source of light, and a screen were placed at I, a spot of light would be formed on the side of the screen facing the mirror provided that the screen were not so large as to prevent all the incident rays from reaching the mirror. Alternatively the image may be seen direct by an eye placed as shown in Fig. 511 so as to receive some of the reflected rays after they have crossed over at I.

If a point object were placed at I it would form a real image at O. The rays forming the image would follow the paths in Fig. 511 but with all directions reversed. The points O and I are said to be **conjugate foci.**

Case 2. There are other possibilities with regard to image formation with a concave mirror. Thus in Fig. 512 the object O is situated nearer the mirror than before. When it is sufficiently near (actually nearer than the midpoint of CP), a typical reflected ray *diverges* from

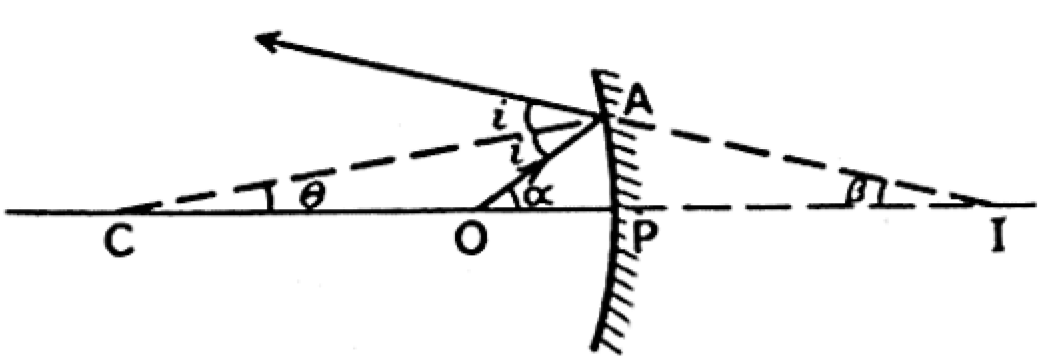


FIG. 512

the axis. Let it be produced backwards and cut the prolongation of the axis at I. In place of the previous relation between the angles α, β and θ we now have

$$\alpha = i + \theta$$

and

$$i = \beta + \theta$$

so that

$$-\beta + \alpha = 2\theta$$

By an exactly similar argument to the previous case this relation between the angles leads to

$$-\frac{1}{\mathrm{IP}} + \frac{1}{\mathrm{OP}} = \frac{2}{\mathrm{CP}}$$

provided that A is always near P, *i.e.* the rays are paraxial. Thus I is the image of O. It is a **virtual** image, and from this point of view is of the same nature as the image formed when a real object is placed in front of a plane mirror. Fig. 513 shows the formation of the image and its observation by a suitably placed eye. It will be noticed that the different character of the image in Case 2 as compared with Case 1 is accompanied by a negative sign in front of the reciprocal of the image distance in the equation.

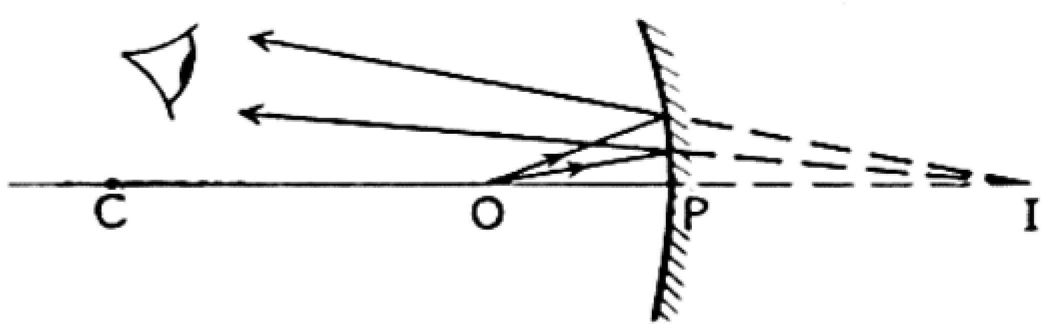


FIG. 513

Case 3. Lastly, we consider the effect of a concave mirror on an incident ray such as that in Fig. 514, *i.e.* one which is converging towards the principal axis and when produced behind the mirror cuts it at the point O. The reflected ray crosses the principal axis at I in front of the mirror. This time the relationship between the angles is

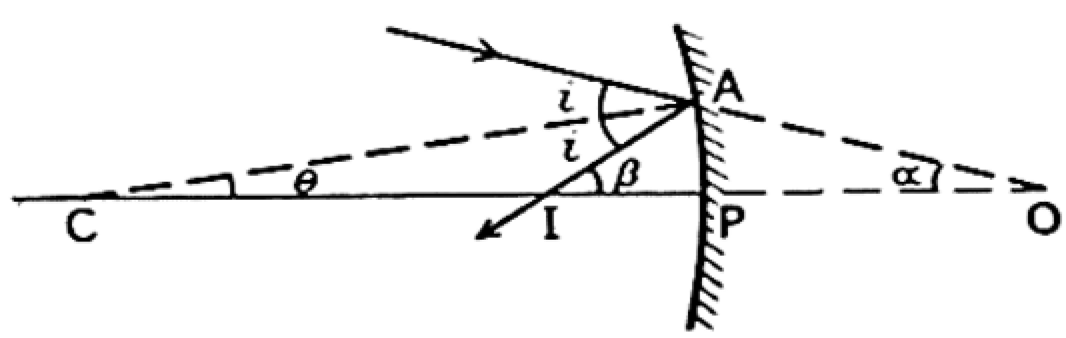


FIG. 514

$$\beta - \alpha = 2\theta$$

which for paraxial rays gives the relation between the distances as

$$\frac{1}{\mathrm{IP}} - \frac{1}{\mathrm{OP}} = \frac{2}{\mathrm{CP}}$$

Thus when a number of paraxial incident rays are converging to a point such as O (a **virtual object**) behind the mirror, a real image is formed in front of the mirror as illustrated in Fig. 515. It will be noticed that if

a point object is placed at I it would form a virtual image at O, the directions of all the rays being reversed. The conditions would then be the same as in Case 2.

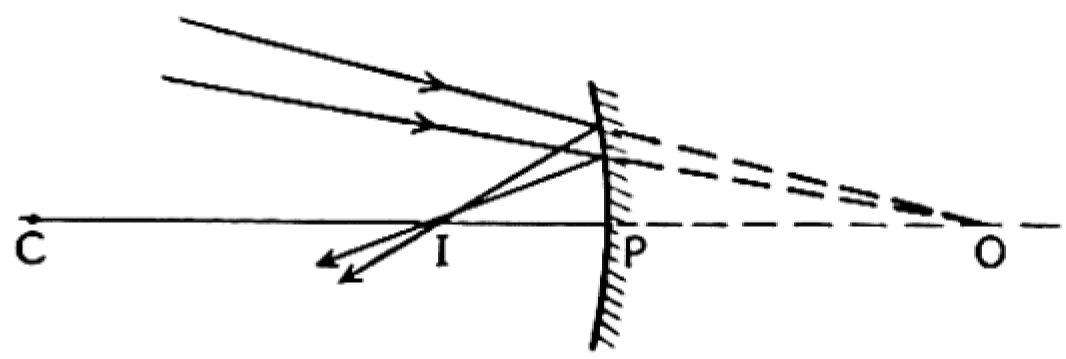


FIG. 515

Case 3 differs from Case 1 in respect of the nature of the object. It is virtual instead of real. This distinction between the two cases is marked by the difference of sign in front of the reciprocal of the object distance.

Correlation of the Three Cases by a Sign Convention.—In each of the three cases of reflection by a concave mirror which have just been discussed, there is the same form of equation connecting the distances of object and image with the radius of curvature of the reflecting surface. The three cases are distinguished by the algebraic signs of the terms involving the object and image distances. It is possible to use one equation for all cases and to modify the signs appropriately for each case according to a definite rule. Such a rule is called a **sign convention.**

There are several sign conventions in use, and their respective merits are frequently debated by teachers. In this book we shall use what is called the **real-is-positive** or **R.P.** convention, not because it is considered to be superior to all others (no convention is without its difficulties), but because it is widely taught and often successfully used by students. We proceed now to explain the meaning and use of this convention.

We have seen that for reflection by a concave mirror there are three different cases. Thus:

Case 1 (object real, image real):

$$\frac{1}{IP}+\frac{1}{OP}=\frac{2}{CP}$$

Case 2 (object real, image virtual):

$$-\frac{1}{IP}+\frac{1}{OP}=\frac{2}{CP}$$

Case 3 (object virtual, image real):

$$\frac{1}{IP}-\frac{1}{OP}=\frac{2}{CP}$$

Since the equations differ from each other only as regards the algebraic signs of the terms, we can combine them into one general equation, namely

$$\frac{1}{v}+\frac{1}{u}=\frac{2}{r} \qquad . \quad . \quad . \quad . \quad . \quad (2)$$

The quantity u in this equation stands not merely for the distance between the object and the pole of the mirror, but for this distance *prefixed with the appropriate algebraic sign* as determined by the convention. Similarly, v has the same significance as regards the image. The quantity r stands for the radius of curvature of the mirror with *its* appropriate sign, which we shall discuss later. We need only state now that the sign of the radius is positive for a concave mirror and negative for a convex mirror.

According to the R.P. sign convention, a positive sign is to be inserted in front of the object distance when the object is real, and a negative sign when the object is virtual. The same rule applies to the sign of the image distance—positive for a real image, negative for a virtual image. Thus if the object is separated from the mirror by l units of length, we must substitute $+l$ for u if the object is real and $-l$ if it is virtual. It will be seen that the use of this convention converts equation (2) into the appropriate equation for any one of the cases 1, 2 and 3.

As an example, suppose we are asked to find the position and nature of the image produced when a virtual object is situated 12 cm. from a concave mirror of 30 cm. radius. In equation (2) we have $u = -12$ cm., $r = +30$ cm. Hence

$$\frac{1}{v} - \frac{1}{12} = \frac{2}{30}$$

$$\therefore \quad \frac{1}{v} = \frac{2}{30} + \frac{1}{12}$$

$$= \frac{3}{20}$$

and

$$v = \frac{20}{3}$$

$$= 6\tfrac{2}{3} \text{ cm.}$$

The sign of this distance is positive so that it refers to a real image, and the answer to the question is that a real image is formed $6\frac{2}{3}$ cm. from the mirror. It is necessarily in front of the mirror if it is real. The conditions are as in Case 3.

It should be noticed carefully that the insertion of the correct signs when we substitute for two of the quantities u, v and r, ensures that the unknown quantity, when calculated, has its appropriate algebraic sign, which we can interpret as part of the information afforded by the calculation. In this connection it must be emphasized that even if, as is often the case, we can forecast by some other means the sign which the unknown quantity will have when the calculation has been made, we should not insert this sign in the equation. The use of the correct

signs for the *known* quantities is sufficient to ensure the correct sign for the unknown quantity.

Focal Length of a Concave Mirror.—Suppose that an object is situated on the principal axis of a concave mirror of radius r at a very great distance from the mirror. Provided that the width of the mirror is small compared with r, we can apply the equation

$$\frac{1}{v}+\frac{1}{u}=\frac{2}{r}$$

and the value of u is ∞ when the object is very distant so that $\frac{1}{u}$ is zero. The equation therefore gives

$$v=\frac{r}{2}$$

Since r is a positive quantity for a concave mirror, v is also positive when $u=\infty$, so that the image is real and is situated midway between the centre of curvature of the mirror and its pole. Another way of describing the state of affairs is to say that the incident rays are parallel to each other and to the principal axis.

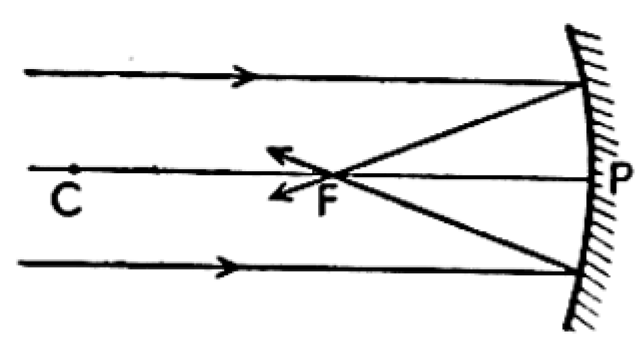


FIG. 516

The position of the image formed by a pencil of parallel rays parallel to the principal axis is called the **principal focus** of the mirror. It is marked F in Fig. 516. The distance FP is called the **focal length** of the mirror and is denoted by f. Thus we have

$$f=\frac{r}{2}$$

and the fundamental equation can now be written

$$\frac{1}{v}+\frac{1}{u}=\frac{2}{r}=\frac{1}{f}$$

Position of Image for Various Positions of the Object.—We now consider the reflected rays produced by a real point object O situated at various positions on the principal axis of a concave mirror. Since the object is real,

$$u=\text{OP and is positive}$$

and for a concave mirror

$$r=\text{CP and is positive}$$

The position of the image (v) is therefore given by

$$\frac{1}{v}+\frac{1}{\mathrm{OP}}=\frac{2}{\mathrm{CP}}$$

or

$$v=\frac{(\mathrm{OP})\times(\mathrm{CP})}{2(\mathrm{OP})-(\mathrm{CP})}$$

$$=\frac{\mathrm{CP}}{2-\dfrac{\mathrm{CP}}{\mathrm{OP}}} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

In the first place, when OP is infinite we have

$$\frac{\mathrm{CP}}{\mathrm{OP}}=0$$

so that, from equation (3),

$$v=\frac{\mathrm{CP}}{2}$$

which simply means that the image is real and is situated at the principal focus F.

Next, when OP is less than ∞ but greater than CP, *i.e.* O is situated between the centre of curvature and infinity, we have

$$\frac{\mathrm{CP}}{\mathrm{OP}}<1$$

so that the value of $\left(2-\dfrac{\mathrm{CP}}{\mathrm{OP}}\right)$ lies between 1 and 2. Therefore, by equation (3), v lies between CP and $\dfrac{\mathrm{CP}}{2}$. Thus the image is real (v is positive) and is situated between F and C (Fig. 517). The points O and I

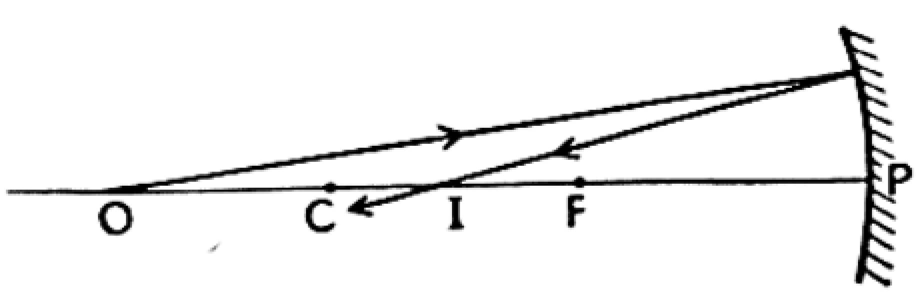


FIG. 517

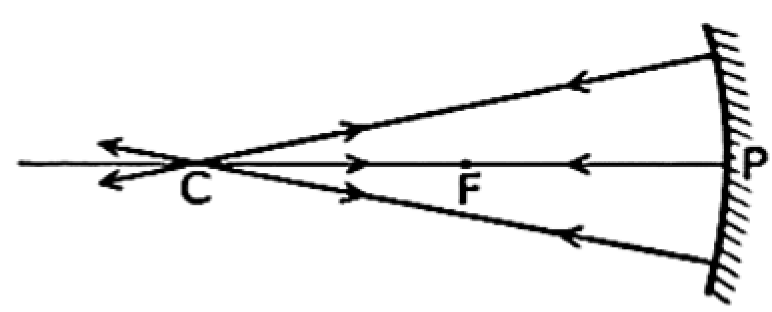


FIG. 518

are conjugate foci (page 711) and can be interchanged, so that when O is situated between F and C, I lies between C and infinity.

When OP is actually equal to CP, *i.e.* the object is at the centre of curvature, v is also equal to CP by equation (3). The incident rays are reflected back along their own paths because they strike the mirror normally (Fig. 518).

When O is situated at F, *i.e.* $\mathrm{OP}=\dfrac{\mathrm{CP}}{2}$, v is equal to ∞. The reflected rays are then all parallel to the principal axis.

Finally, when O lies between P and F, we have

$$OP < \frac{CP}{2}$$

so that

$$\frac{CP}{OP} > 2$$

Therefore $\left(2 - \frac{CP}{OP}\right)$ is negative, v is negative by equation (3) and the image is virtual. This occurs when a real object is nearer to the mirror than the principal focus.

Convex Mirror.—We investigate the action of a convex mirror on an incident pencil of rays in exactly the same way as we did with the concave mirror.

Case 1. Let a point object be situated at O (Fig. 519) on the principal axis. An incident ray OA is reflected as shown according to the laws of reflection, CA produced being the normal. The three angles marked i are equal, $\widehat{IAC}$ being vertically opposite to the angle of reflection. The reflected ray is diverging from the principal axis, and when it is produced backwards it cuts the axis at I. The relationship between the angles marked is

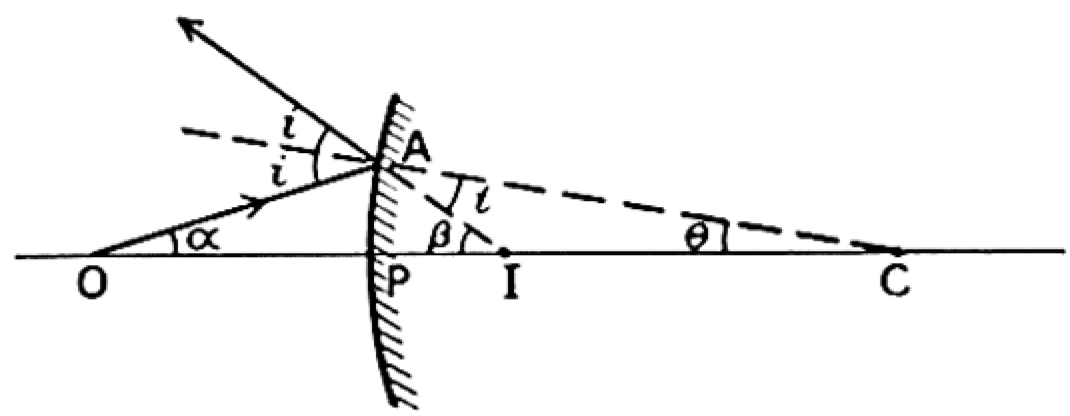


FIG. 519

$$\beta = i + \theta \text{ in triangle AIC}$$

and

$$\theta = i - \alpha \text{ in triangle ACO}$$

Eliminating i we get

$$\beta - \alpha = 2\theta$$

Assuming that the rays are paraxial this leads, in exactly the same way as for the concave mirror, to

$$\frac{1}{IP} - \frac{1}{OP} = \frac{2}{CP}$$

Thus all paraxial rays from O will, after reflection, appear to diverge from I, which is the *virtual* image of O (Fig. 520). This is the only type of

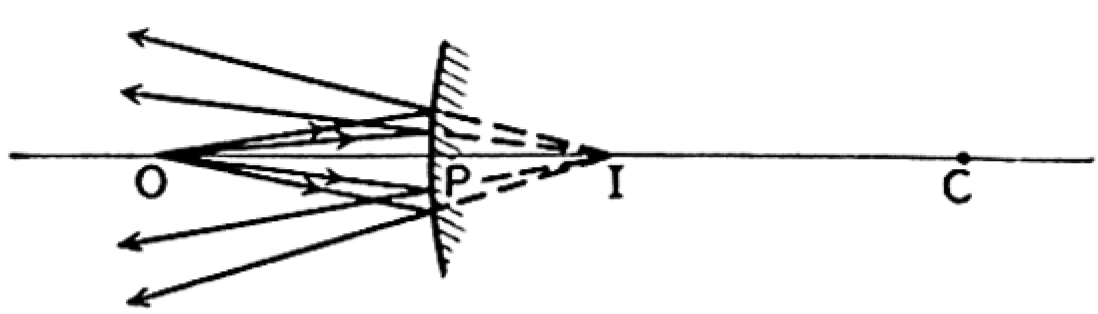


FIG. 520

reflection obtained with a convex mirror when the object is real. Two other types are possible with a virtual object, and these are now dealt with.

Case 2. Let an incident ray be travelling towards O (Fig. 521). When OP is greater than CP/2, the reflected ray diverges from the axis and cuts

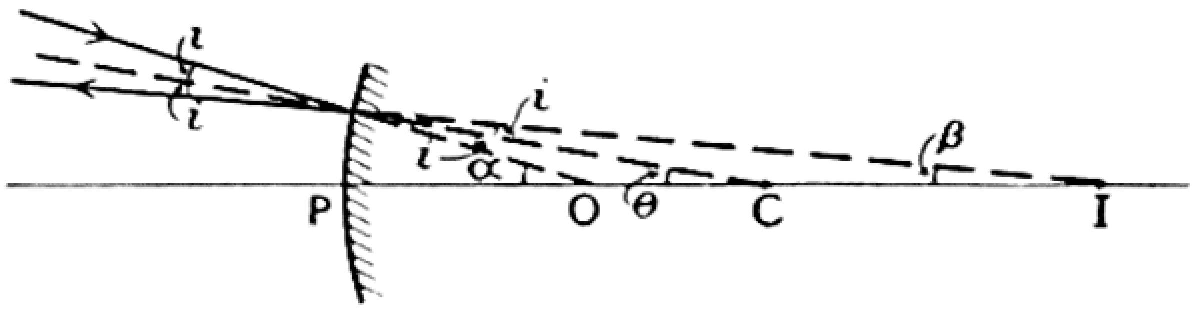


FIG. 521

it at I when produced backwards. By the same method as for previous cases it can be shown that for paraxial rays

$$\frac{1}{IP}+\frac{1}{OP}=\frac{2}{CP}$$

Thus I is the virtual image of the virtual object O (Fig. 522). The two points are conjugate foci and they can be interchanged by reversing the directions of the light rays.

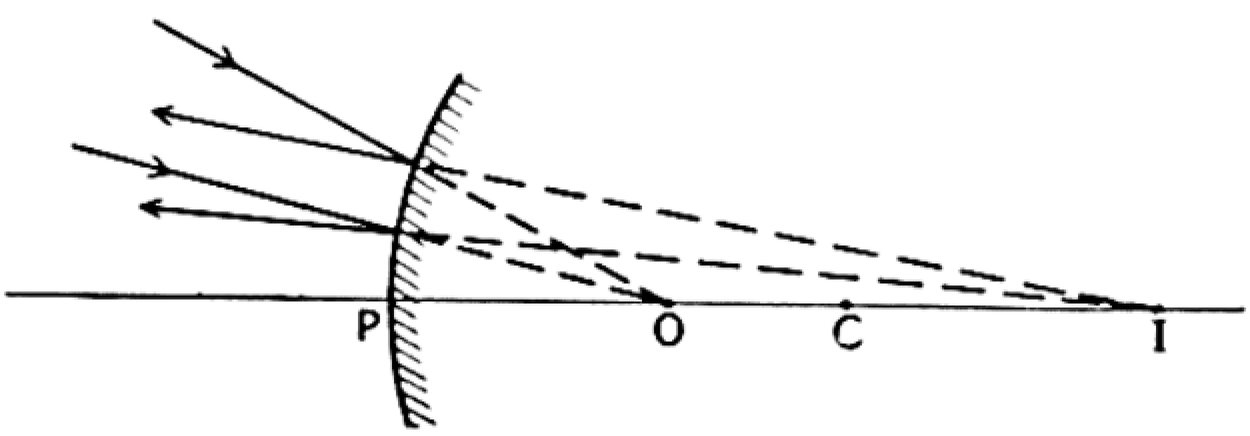


FIG. 522

Case 3. Finally, let an incident ray be directed at a point O which is sufficiently near the mirror (actually nearer than the midpoint of CP) to cause the reflected ray to cut the axis at I in front of the mirror (Fig. 523). The

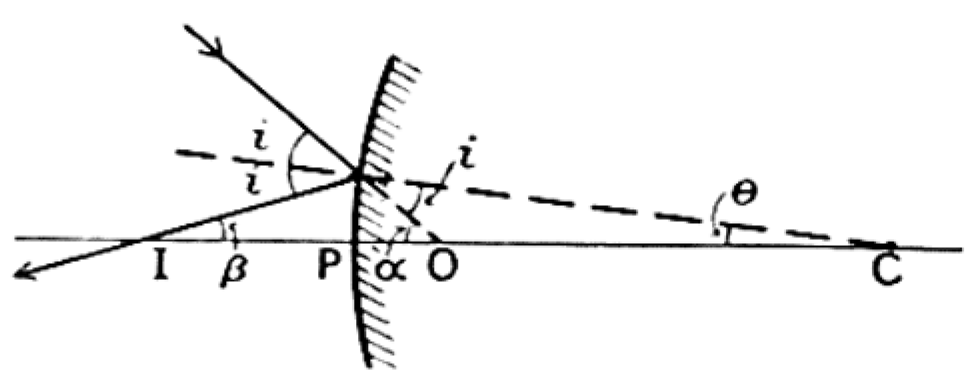


FIG. 523

equation for paraxial rays in this case can be derived from first principles as

$$-\frac{1}{IP}+\frac{1}{OP}=\frac{2}{CP}$$

The object O is virtual and the image is real (Fig. 524). This is the exact reverse of Case 1.

Use of the Sign Convention.—The three types of image formation by a convex mirror can all be incorporated in the equation

$$\frac{1}{v}+\frac{1}{u}=\frac{2}{r}$$

where u and v have exactly the same significance as with the concave mirror and are subject to the same sign convention, and r stands for $-\mathrm{CP}$, *i.e.* the radius of curvature of the mirror with a negative sign prefixed. With the concave mirror, it will be remembered, r was equal to CP (positive sign). We must note, therefore, that our fundamental equation and sign convention require that the radius of curvature of a concave mirror is positive and that of a convex mirror is negative.

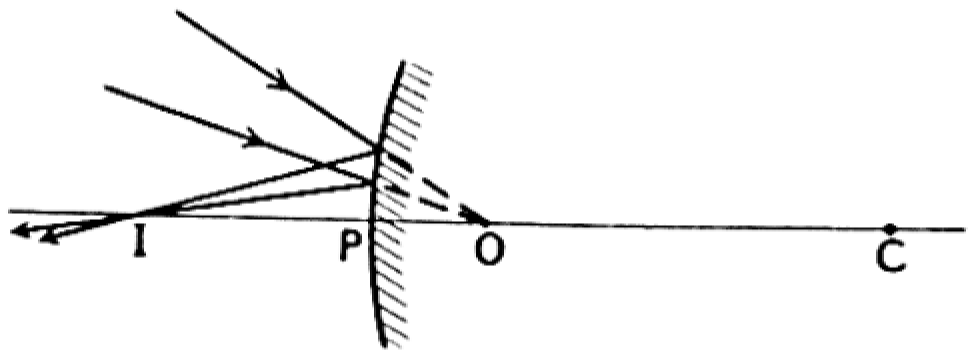


FIG. 524

Focal Length of a Convex Mirror.—When a real object is situated at an infinite distance from the mirror, u is equal to ∞, $\frac{1}{u}$ is zero, and therefore

$$\frac{1}{v}=\frac{2}{r}$$

$$=-\frac{2}{\mathrm{CP}},\quad \text{since } r=-\mathrm{CP}$$

and

$$v=-\frac{\mathrm{CP}}{2}$$

The negative sign indicates that the image is virtual. It is situated midway between P and C at F (Fig. 525), which is the **principal focus** of the mirror. The focal length f is given by

$$f=-\frac{\mathrm{CP}}{2}$$

It is clear that in both the concave and the convex mirrors f has the same

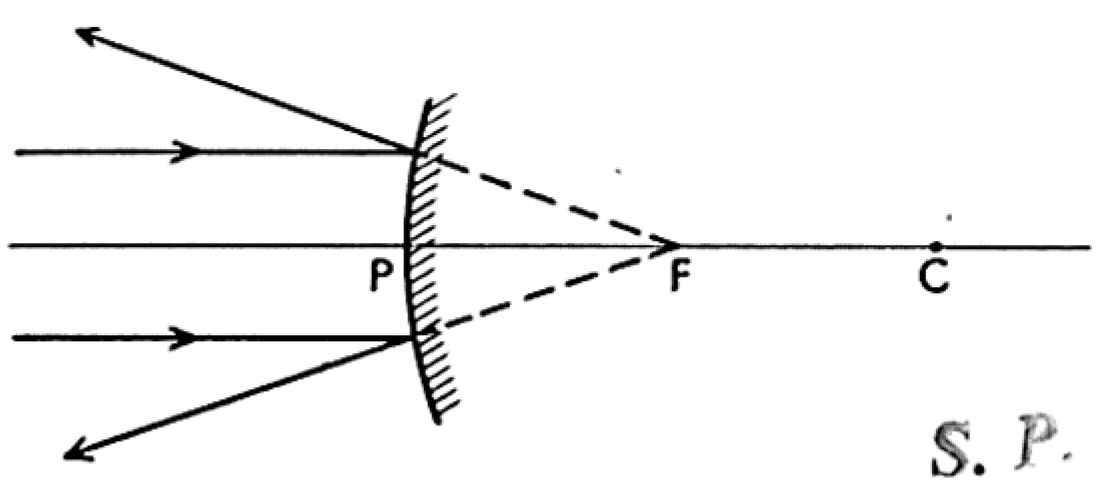


FIG. 525

sign as r. Both f and r are positive for a concave mirror and negative for a convex mirror.

When a real object is placed anywhere on the principal axis in front

of a convex mirror the appropriate equation is derived from

$$\frac{1}{v}+\frac{1}{u}=\frac{2}{r}=\frac{1}{f}$$

where

$$u=\text{OP and is positive}$$

and

$$r=2f=-\text{CP and is negative}$$

In a manner similar to that used on page 716 for a concave mirror, the substitution leads to the following expression for v:—

$$v=-\frac{\text{CP}}{2+\dfrac{\text{CP}}{\text{OP}}}$$

When the object is at infinity (OP $=\infty$), v has its maximum magnitude and is equal to $-\frac{\text{CP}}{2}$, *i.e.* the image is at the principal focus, as already explained. As OP is decreased from ∞ to 0, the magnitude of v decreases from $\frac{\text{CP}}{2}$ to 0, its sign being always negative. Thus the image of a real object in a convex mirror is always virtual and its position moves from F to P as the object moves from infinity to P.

Image Due to Two Mirrors.—It is often necessary to work out the position of the final image when light is reflected first in one mirror and then in a second mirror. For mirror (1) the relation between the object and image distances is

$$\frac{1}{v_1}+\frac{1}{u_1}=\frac{2}{r_1}$$

and the image I_1 formed by mirror (1) acts as the object for mirror (2). There are three different possibilities as to the relative positions of the image I_1 and the two mirrors. In Case (i) (Fig. 526 (i)) the mirror (1) has formed a real image at I_1 between the two mirrors which are shown as dotted lines, since the present argument is independent of the natures of the mirrors. The fronts of the mirrors must, of course, be facing each other. If P_1 and P_2 are the poles of the mirrors, we have, since I_1 is a real image,

$$v_1=+I_1P_1$$

It is evident from the drawing of the rays by which I_1 is formed that I_1 acts as a real object for mirror (2). Therefore, in the equation for reflection at mirror (2), namely

$$\frac{1}{v_2}+\frac{1}{u_2}=\frac{2}{r_2}$$

we have

$$u_2 = +I_1P_2$$

If d is the distance between the mirrors (always considered as a positive quantity), it is evident from the diagram that

$$I_1P_1 + I_1P_2 = d$$

i.e., substituting for I_1P_1 and I_1P_2,

$$v_1 + u_2 = d$$

In Case (ii) (Fig. 526 (ii)) I_1 is a virtual image behind mirror (1), so that

$$v_1 = -I_1P_1$$

It is important to realize that in this case also I_1 acts as a real object for

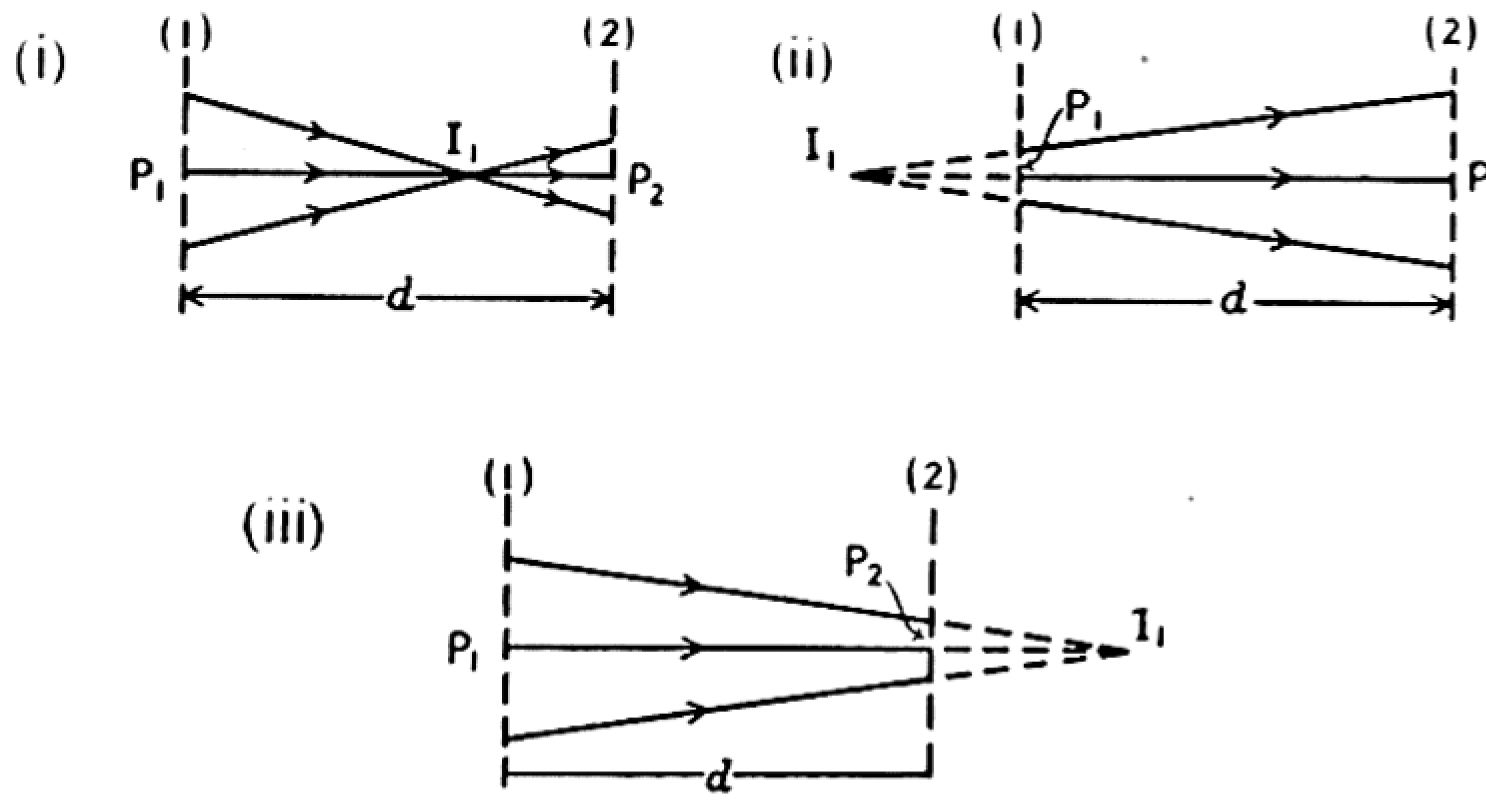


FIG. 526

mirror (2), because the rays strike mirror (2) in just the same way as if a real object were situated at I_1. For mirror (2), therefore,

$$u_2 = +I_1P_2$$

Evidently

$$I_1P_2 - I_1P_1 = d$$

so that

$$u_2 + v_1 = d$$

which is the same relationship as for Case (i). Finally, in Case (iii) (Fig. 526 (iii)), I_1 would have been a real image formed by mirror (1) but for the presence of mirror (2), which intercepts the rays before they reach their focus at I_1. Therefore

$$v_1 = +I_1P_1$$

The image I_1 acts as a virtual object for mirror (2), so that

$$u_2 = -I_1P_2$$

Now
$$I_1P_1 - I_1P_2 = d$$
so that
$$v_1 + u_2 = d$$

which is again the same equation. This relationship can therefore be used for all cases. It works automatically, like the mirror equation, when the sign convention is used.

***Example.**—Two spherical mirrors, (1) and (2), are placed facing each other 10 cm. apart. Mirror (2) has a radius of curvature of 40 cm. and is convex. An object, placed 78 cm. in front of mirror (1), forms, by reflection first in (1) and then in (2), a real image 15 cm. in front of (2). Find the nature and radius of curvature of mirror (1), and draw a pencil of rays from the object to the final image.*

In Fig. 527 (i) the object is represented by O and the final image by I_2. It is necessary of course that the mirrors should be slightly inclined to each other in

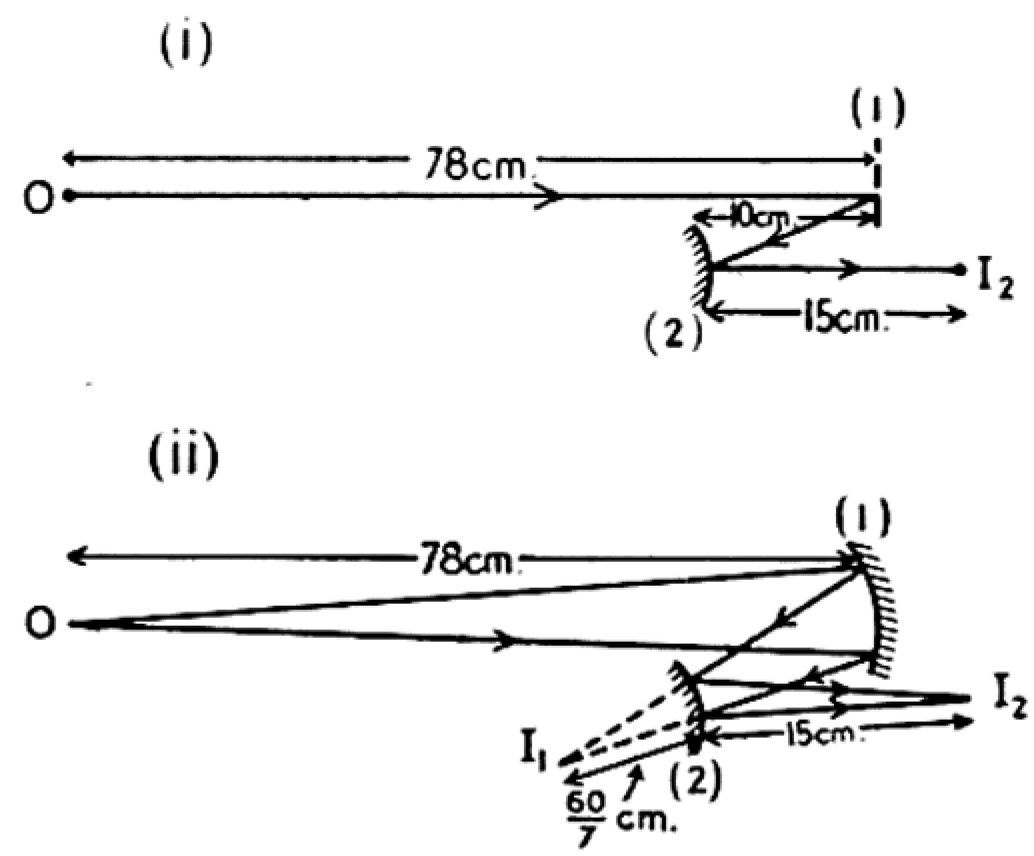


FIG. 527

order that (2) shall not obstruct the light from (1), but we can suppose that the mirrors are small so that their inclinations are small and the ordinary equations for paraxial rays can be used.

We have to find the magnitude and sign of r for mirror (1). Let it be r_1. For the reflection at mirror (1) we can write

$$\frac{1}{v_1} + \frac{1}{u_1} = \frac{2}{r_1}$$

where $u_1 = 78$ cm. (positive sign because O is a real object), and v_1 refers to the image formed by (1) and is unknown. We can, however, find v_1 by working backwards from the final image. Thus for the reflection in mirror (2),

$$\frac{1}{v_2} + \frac{1}{u_2} = \frac{2}{r_2}$$

where
$$v_2 = 15 \text{ cm.} \qquad \text{(real image)}$$
and
$$r_2 = -40 \text{ cm.} \qquad \text{(convex mirror)}$$

Substituting, we have

$$\frac{1}{15}+\frac{1}{u_2}=-\frac{2}{40}$$

whence

$$u_2=-\frac{60}{7}\text{ cm.}$$

We may note in passing and as a guide to drawing the ray diagram that the negative sign of u_2 implies that mirror (2) has a virtual object. Since we have shown that the equation

$$v_1+u_2=d$$

(where d is the separation of the mirrors) applies to all possible cases of the type involved in this problem, we can immediately find v_1 by inserting the values of u_2 and $d(=10$ cm.). Thus

$$v_1-\frac{60}{7}=10$$

so that

$$v_1=\frac{130}{7}\text{ cm.}$$

We now substitute this value for v_1 in the equation for the first reflection and obtain

$$\frac{7}{130}+\frac{1}{78}=\frac{2}{r_1}$$

or

$$\frac{2}{r_1}=\frac{21+5}{390}$$

$$=\frac{26}{390}$$

Whence

$$r_1=30\text{ cm.}$$

Therefore the radius of mirror (1) is 30 cm. and the positive sign of r_1 indicates that the mirror is *concave*. The pencil of rays from the object to the final image is shown in Fig. 527 (ii), in which I_1 is the first image and I_2 the final image.

2. IMAGES OF FINITE OBJECTS

We now consider what happens when the object is not a single point but a short line such as the arrow OO′ (Fig. 528) lying perpendicular to the principal axis.

The foot of the arrow (O) lies on the principal axis, and its image I will also lie on this line. Now let the line representing the principal axis be rotated about the centre of curvature C until it reaches the position O′CP′. Although, according to the definition which we have adopted on page 709, this line is not a principal axis because it does not pass through the central point P of the mirror, yet, as regards the relative positions of point objects and their images situated upon it, it has the same

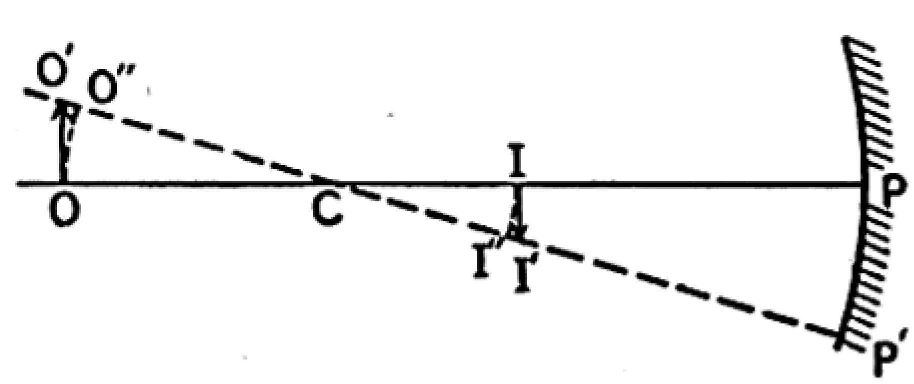


FIG. 528

properties as the principal axis. The fact that P′ is not the central point of the mirror does not prevent the usual mirror equation from being valid for point objects on the line through P′ and C, the distances of object and image being measured along the line to the point P′. On rotating the line from OCP to O′CP′ the point O will describe a short circular arc OO″ about C, while its image, retaining its position on the rotating line, will move along the circular arc II″, also about C. If the size of the finite object OO′ is small the rotation of the axis is small, and in these circumstances O′ and O″ are very close together and so are I′ and I″. Therefore since I″ is the image of O″, we can regard I′ as the image of O′.

It follows, therefore, that in order to find the image of the arrow OO′ it is only necessary to find I, the image of O, on the principal axis by calculation. We then draw II′ perpendicular to the axis to cut O′C produced at I′, and II′ is the required image. It can be seen immediately that the image is inverted with respect to the object.

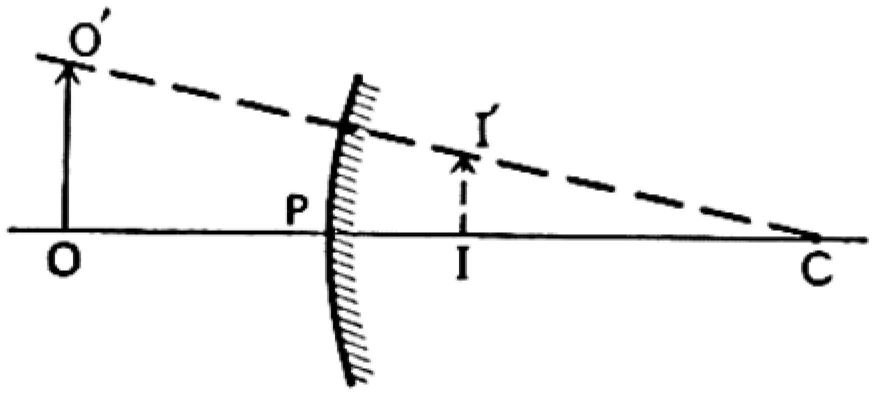


FIG. 529

This principle can be applied to all the possible cases of reflection by concave and convex mirrors although we have only established it for one case. Fig. 529 shows its application to a convex mirror with a real object. The image in this case is upright and virtual.

When the object is very far from a small mirror we can regard the light rays striking the mirror from *any one point* on the object as being parallel to each other (Fig. 530 (i) for a concave mirror and (ii) for a convex

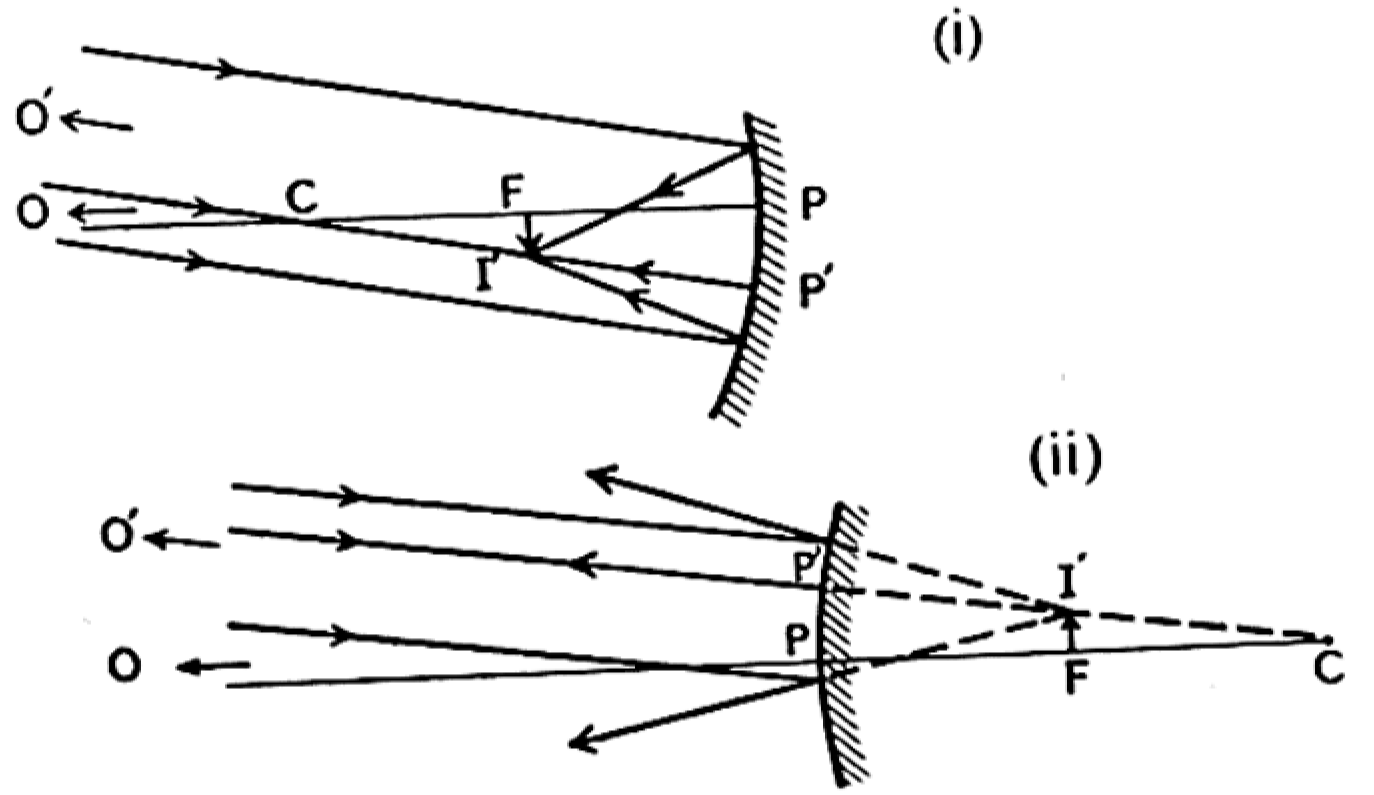


FIG. 530

mirror). But unless the object is as small as the mirror the rays coming from O′ will not be parallel to those from O. Thus O will form an image at the principal focus F since the rays coming from it are parallel to the

principal axis, while the parallel rays coming from O′ form an image at I′ which lies on the axis CP′ parallel to the incident light, the distance I′P′ being equal to PF. By the same principle as has been used in the previous cases we can assume that the image FI′ is perpendicular to the principal axis. Thus a distant object forms an image lying in a plane through the principal focus perpendicular to the principal axis. This is called the **focal plane** of the mirror.

If any of the diagrams of finite objects and images are rotated about the principal axis, the lines OO′ and II′ describe plane circles, while the line representing the mirror describes the actual surface of the mirror. In this way we extend our investigation to a plane two-dimensional object, and it is evident that it forms a plane image at a distance from P which is related to the object distance by the same equation as for point objects and images.

Magnification of the Image.—In the foregoing diagrams the fact that I′ is the image of O′ means that all *paraxial* rays striking the mirror from O′ are reflected to pass through I′. In particular the incident ray from O′ to the pole of the mirror (Fig. 531 (i) and (ii)) is reflected so as to make the two angles i equal. The triangles OPO′ and IPI′ are similar because of the equality of the angles i and of the right angles at the feet of the arrows. Therefore in each case we have

$$\frac{\mathrm{II'}}{\mathrm{OO'}}=\frac{\mathrm{IP}}{\mathrm{OP}}$$

The same relation is true for all the possible cases of reflection by a spherical mirror.

The **transverse linear magnification** is defined as the ratio of the length of a line on the image perpendicular to the axis on which the object and image lie to the length of the corresponding line on the object.

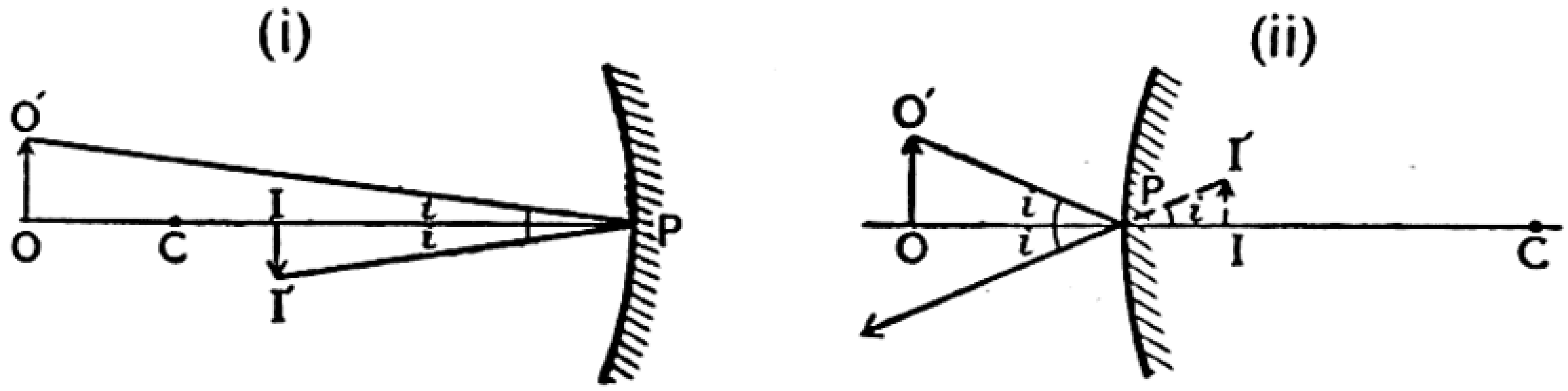


Fig. 531

We have already expressed this quantity in terms of the distances IP and OP, but it is convenient to go a step further and to relate it with u and v. In the case shown in Fig. 531 (i) the object and image are both real, and we have

$$\frac{v}{u}=\frac{\mathrm{IP}}{\mathrm{OP}}=\frac{\mathrm{II'}}{\mathrm{OO'}}=\text{magnification}$$

In Fig. 531 (ii), on the other hand, the image is virtual so that v is equal to $-$IP, and we have

$$\frac{v}{u}=\frac{-\mathrm{IP}}{\mathrm{OP}}=-\frac{\mathrm{II'}}{\mathrm{OO'}}=-(\text{magnification})$$

If we define a quantity m by

$$-m=\frac{v}{u}$$

then m represents the magnification with an algebraic sign prefixed which is positive when the object and image are the same way up and negative when there is inversion. In using the relation $-m=v/u$ in order, say, to find v when the object distance and the magnification are known, we must give each of the known quantities its correct sign according to the convention we have adopted. Thus if the object is given as real and situated 10 cm. from the mirror, and the image is given as virtual with a transverse magnification of 3, then we know that the image is upright with respect to the object so that $m=+3$. Since $u=+10$ cm. this gives $v=-30$ cm. The following example illustrates the matter.

Example.—*A real object placed* 20 *cm. in front of a spherical mirror forms an image with a transverse linear magnification of* 3. *Find the nature and radius of curvature of the mirror when the image is* (a) *virtual*, (b) *real.*

The equations are

$$\frac{1}{v}+\frac{1}{u}=\frac{2}{r}$$

$$-m=\frac{v}{u}$$

together with the R.P. sign convention.

In (*a*) we have

$$u=20 \text{ cm.}$$

and, since the image is virtual, it is the same way up as the object, so that

$$m=+3$$

Therefore

$$v=-mu$$
$$=-60 \text{ cm.}$$

Hence

$$-\frac{1}{60}+\frac{1}{20}=\frac{2}{r}$$

or

$$\frac{r}{2}=\frac{60}{2}$$

so that

$$r=60 \text{ cm.}$$

The mirror therefore has a radius of 60 cm. and is concave (positive sign for r).

In (*b*) we have

$$u=20$$

and, as the image and object are both real, they are inverted with respect to each

other, so that

$$m = -3$$

Therefore

$$v = -mu$$

$$=60 \text{ cm.}$$

Hence

$$\frac{1}{60} + \frac{1}{20} = \frac{2}{r}$$

or

$$\frac{r}{2} = \frac{60}{4}$$

so that

$$r = 30 \text{ cm.}$$

Therefore the mirror has a radius of 30 cm. and is concave.

3. GRAPHICAL CONSTRUCTIONS

General Principles.—It should be remembered that by "principal axis" we mean the straight line through the central point (pole) of the mirror and the centre of curvature. When a point object is situated off the principal axis of a spherical mirror it is possible to find the position of its image by means of graphical construction or ray-tracing according to certain principles which we have already established. These are re-stated below in the form appropriate to the construction. Corresponding statements for the concave and convex mirrors are really identical, but they are written down separately for the sake of clarity in wording.

Concave Mirror

1. An incident ray which passes through the centre of curvature C strikes the mirror normally and is reflected back along its own path.
2. An incident ray parallel to the principal axis is reflected so as to pass through the principal focus F.
3. An incident ray passing through the principal focus F is reflected parallel to the principal axis.

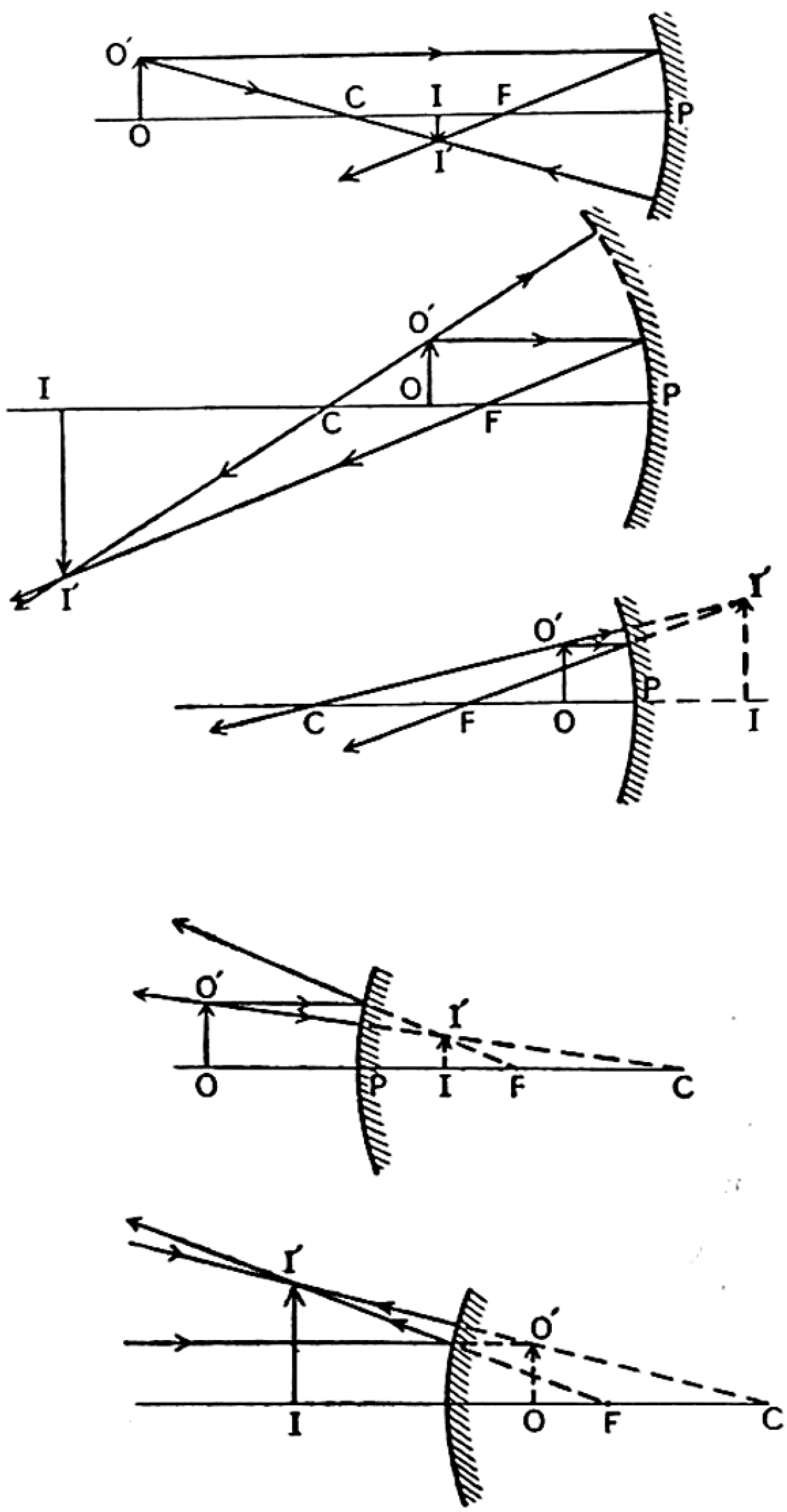
Convex Mirror

1. An incident ray travelling towards C (which is behind the mirror) is reflected back along its own path.
2. An incident ray parallel to the principal axis is reflected so as to appear to come from F (which is behind the mirror).
3. An incident ray travelling towards F (behind the mirror) is reflected parallel to the principal axis.

Suppose that a finite object OO′ is situated in front of a mirror and is at right angles to the principal axis. If O is on the axis and O′ is off it, the position of the image (I′) of the latter point can be found graphically. It

has already been shown (page 724) that, for small objects, when the object is perpendicular to the axis the image is also, so that the image of OO′ is a line drawn perpendicular to the axis from I′.

Examples of Graphical Construction.—It is very helpful to verify the solution of a problem by the method of ray-tracing, and this should always be done where possible. In Fig. 532 (i) to (v) the graphical construction is made for various cases, and notes as to the position and nature of the image are added.



Fɪɢ. 532

(i) *Concave mirror*
OO′ (real) beyond C. II′ inverted, real, diminished, between F and C.

(ii) *Concave mirror*
OO′ (real) between F and C. II′ inverted, real, magnified, beyond C.

(iii) *Concave mirror*
OO′ (real) between P and F. II′ upright, virtual, magnified, behind mirror.

(iv) *Convex mirror*
OO′ (real) situated anywhere in front of mirror. II′ upright, virtual, diminished, between F and P.

(v) *Convex mirror*
OO′ (virtual) between F and mirror. II′ upright, real, magnified, in front of mirror.

4. EXPERIMENTS WITH SPHERICAL MIRRORS

The Optical Bench.—In order to facilitate the setting up of the apparatus and the measurement of distances, the various components necessary in experiments on spherical mirrors and lenses are usually mounted on stands which slide easily along a rail known as an "optical bench." One pattern of bench is illustrated in Fig. 533.

A luminous object is usually used for the formation of real images.

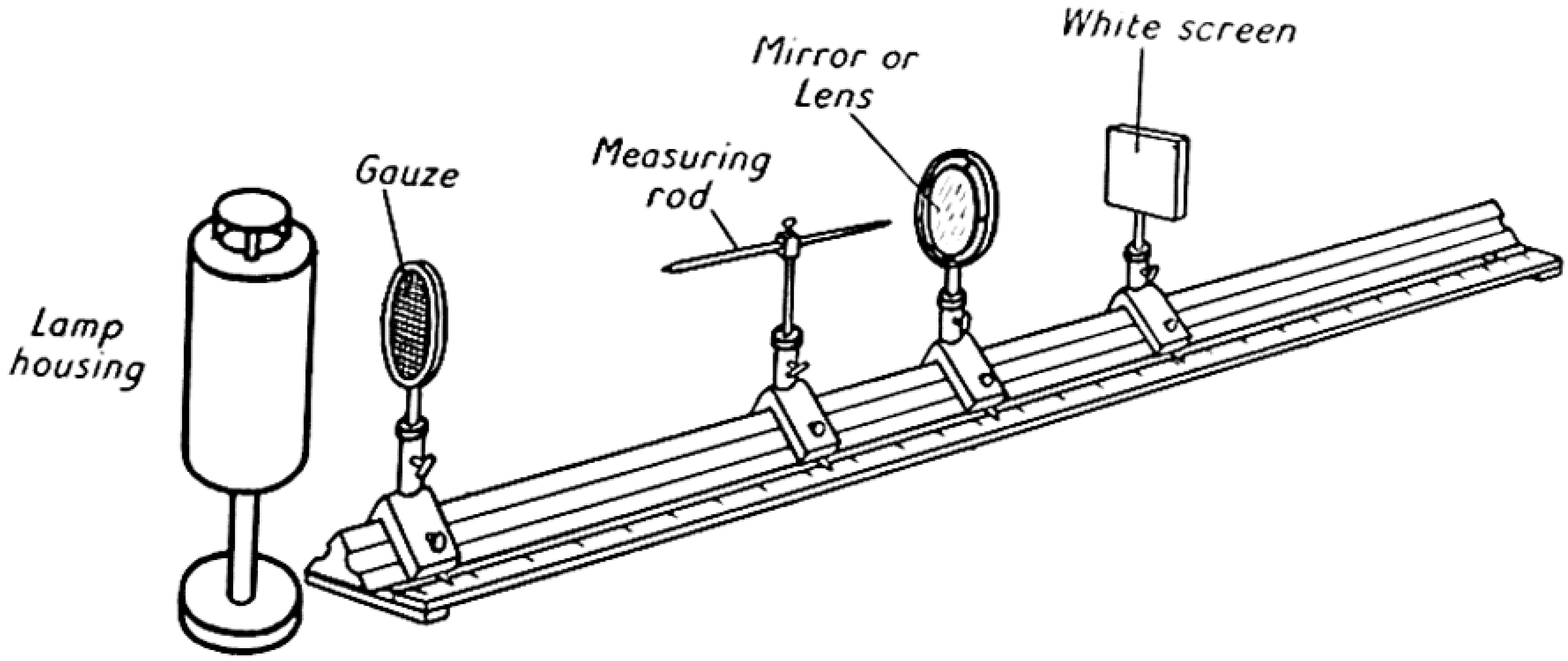


FIG. 533

This frequently consists of a piece of wire gauze or a single pair of cross-wires placed across a hole in a metal plate. The back of the hole is covered with a sheet of ground glass, and a small electric lamp is placed behind this as shown in Fig. 533 or 534. For experiments with lenses it is necessary to use light of one colour (see page 834).

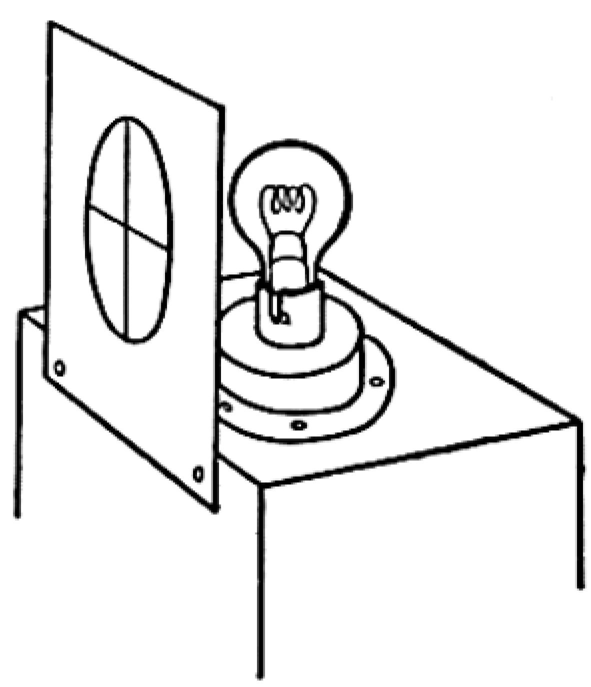


FIG. 534

Each stand on the optical bench has a pointer or vernier on it which moves over a millimetre scale fixed parallel to the rail. Distances between the various components, *e.g.* the mirror and the object, are *not* measured by taking the readings of the marks on their stands and subtracting, because the components in question may not be vertically above their marks. Instead, a metal rod, say 15 cm. long, is fixed horizontally in an additional sliding stand. Its use is shown in Fig. 535 (i) and (ii). In (i) the distance between the two components A and B is equal to the difference between the readings of

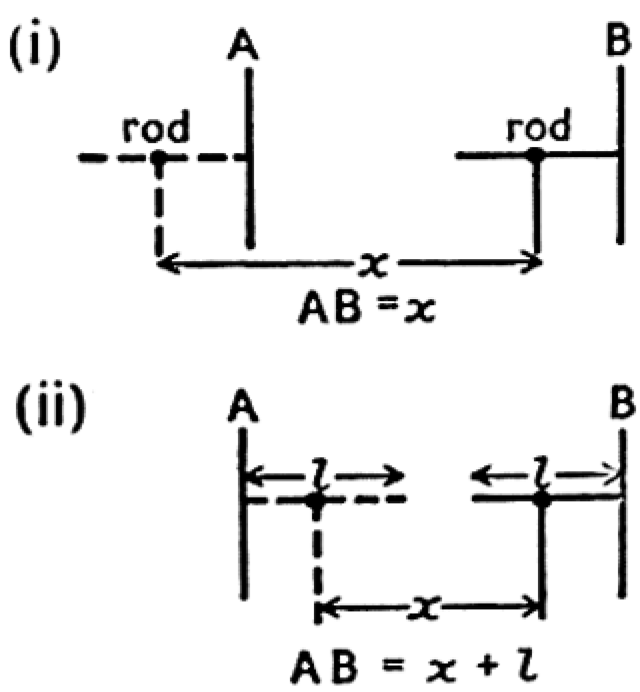


FIG. 535

the mark on the stand when one end of the rod is touching first A and then B. Alternatively (Fig. 535 (ii)) the rod may be placed between the two components and moved so that one of its ends touches A. The reading of its mark is then taken, and it is shifted along the bench until its other end touches B and the reading again taken. The distance between A and B is then equal to the difference of the readings *plus* the length of the rod.

Location of Images by the Elimination of Parallax.—We usually aim at forming real images during optical measurements because their positions can easily be located by moving a white screen about until the image is sharpest. This method is not available for virtual images, however, and another principle must be used.

Suppose that we are looking into a convex mirror M (Fig. 536 (i) and (ii)) and observing the image I of, say, a large vertical pin P_1 placed in front

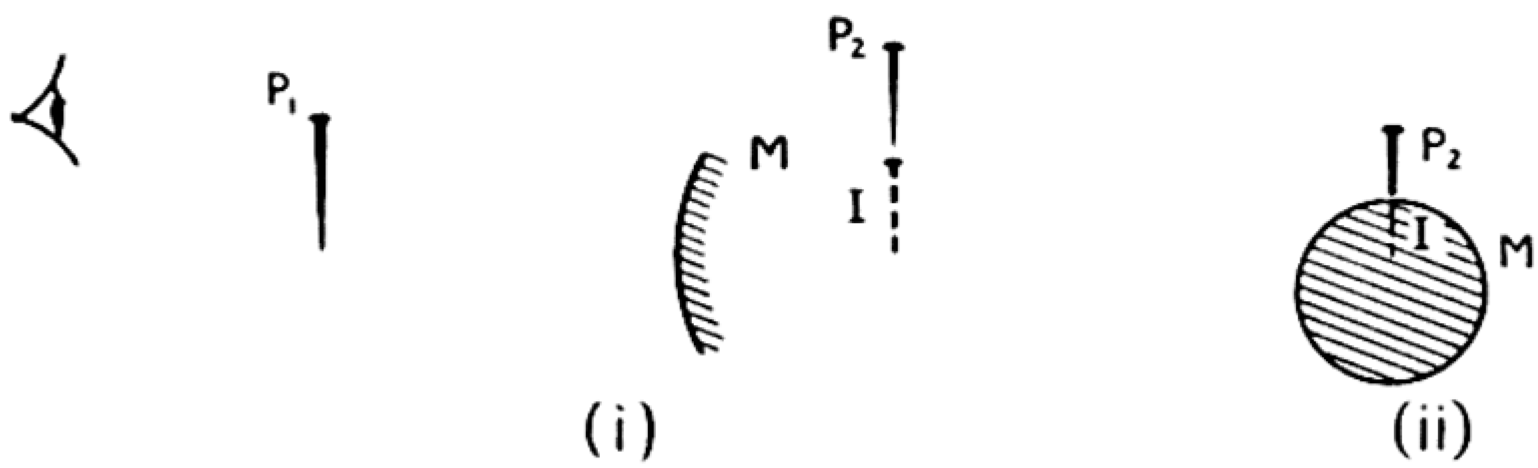


FIG. 536

of it. The image is, of course, behind the mirror. A second pin, P_2, is held behind the mirror, but raised up sufficiently to enable the top part of it to be seen over the top of the mirror. The eye is then moved until I seen in the mirror and P_2 seen over the top are continuous, as shown in Fig. 536 (ii). On moving the eye from left to right and back again, I and P_2 will, in general, be seen to separate. Whichever of the two moves with respect to the other in the *same* direction as the eye is further from the eye—a fact which can easily be verified by experimenting with two real objects. The pin P_2 is therefore moved to or from the mirror until it appears continuous with I wherever the eye is placed so as to see both simultaneously. There is then *no parallax* between P_2 and I and they are coincident in space. Thus I has been located by P_2 and its distance from the mirror can be measured. A real image (I, Fig. 537) situated in front of a mirror can be located similarly and, in fact, rather more easily.

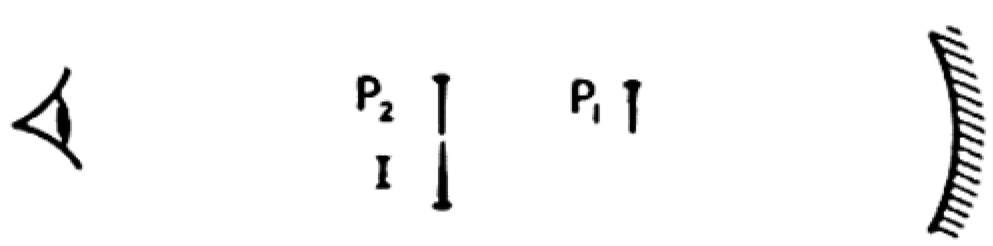


FIG. 537

Determination of *f* and *r* for a Concave Mirror.—The following experiments are described in principle only, because experimental details must be learnt in the laboratory.

Method 1. *Distant Object.*—The image of a fairly bright distant object, *e.g.* a window or clouds, is formed on a screen S (Fig. 538). The distance from the mirror to the screen when the image I is clearest is equal to f. It will be appreciated that whenever a real image due to a mirror is formed on a screen, the latter is bound to obstruct the incident light to some extent. The apparatus is always arranged, however, so that the object and image are as near to the principal axis as possible. In the above experiment and in all others in which real images are formed the image can be located by the elimination of parallax.

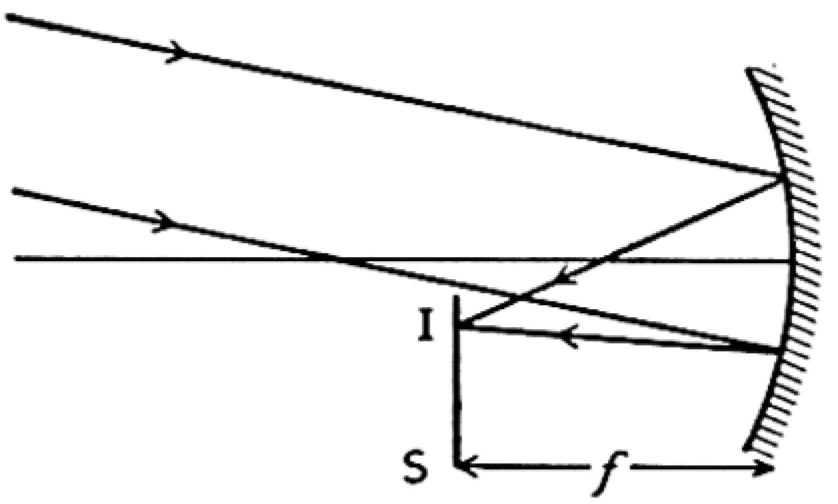


FIG. 538

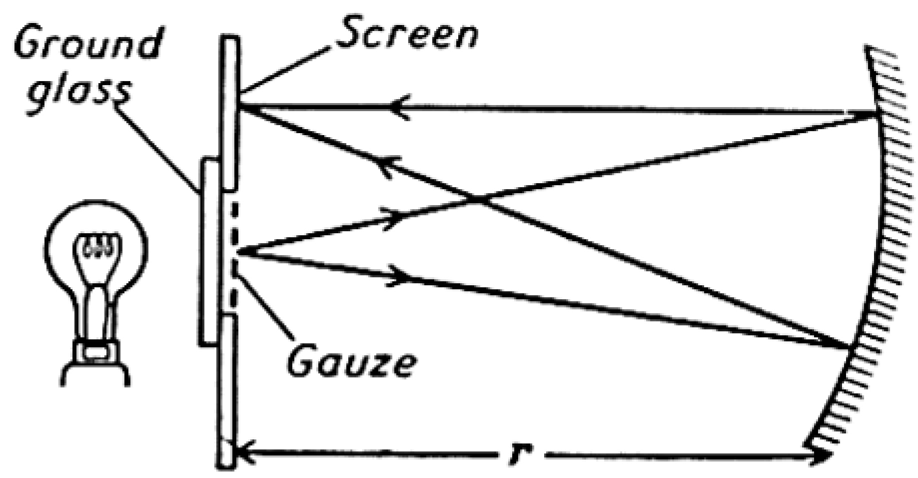


FIG. 539

Method 2. *Direct Determination of r.*—The luminous object such as that shown in Fig. 534 is moved to and from the mirror until an image of the gauze is formed on the front of the metal screen surrounding it. The image and object are now equidistant from the mirror and their distance must be equal to r, which can easily be measured (Fig. 539). The coincidence of a pin with its own image can easily be arranged by the parallax method.

Method 3. *Measurement of u and v.*—The real image of an object is located by means of a screen or by the elimination of parallax, and the distances of object and image are measured for a variety of sensibly chosen object positions. It is also possible, but usually not so accurate, to include some values of u and v for virtual images located by the parallax method. There are several ways of deducing f and r from the readings, a selection of which are mentioned here.

(*a*) Substitute corresponding values of u and v in the mirror equation, taking account of signs, and calculate f from each pair. Then take the mean of the results.

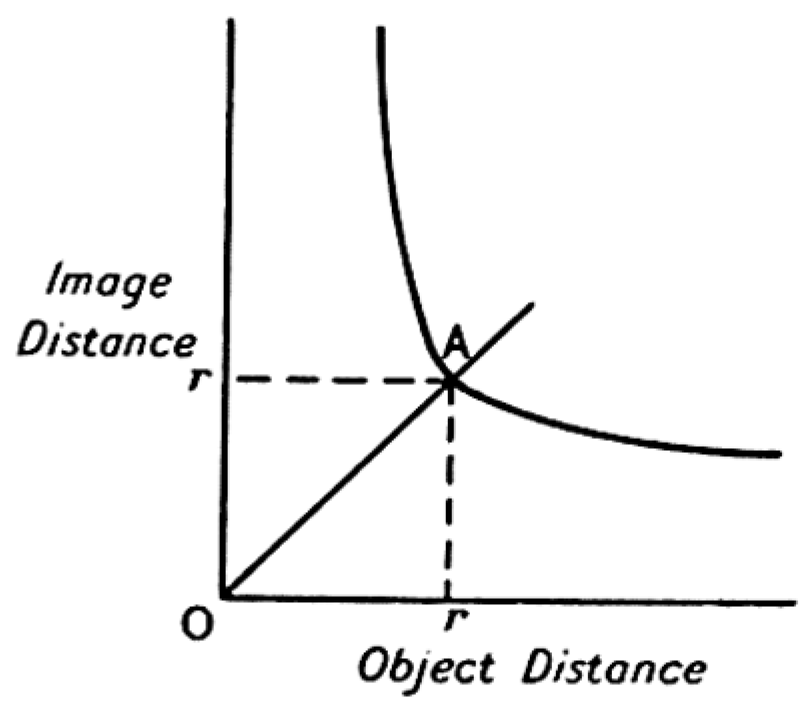


FIG. 540

(*b*) Draw a graph of image distance against object distance for real images, using the same scale on each axis. The shape of this is shown in Fig. 540. Draw a line OA through the

origin, making equal angles with the axes. At the point A where this line crosses the curve the object and image distances are each equal to r.

(c) Draw a graph of $\frac{1}{v}$ against $\frac{1}{u}$. The result (Fig. 541) is a straight line which has a slope of -1 and an intercept of $\frac{1}{f}$ on each axis. This is clear from the equation

$$\frac{1}{v} = -\frac{1}{u} + \frac{1}{f}$$

Since we have used a real object $\frac{1}{u}$ is positive, and the graph should be plotted as shown. The portion of the graph between A and B refers to real objects and images. Its extension across the axis of $\frac{1}{u}$ refers to virtual images, while the object is virtual on the left of the axis of $\frac{1}{v}$.

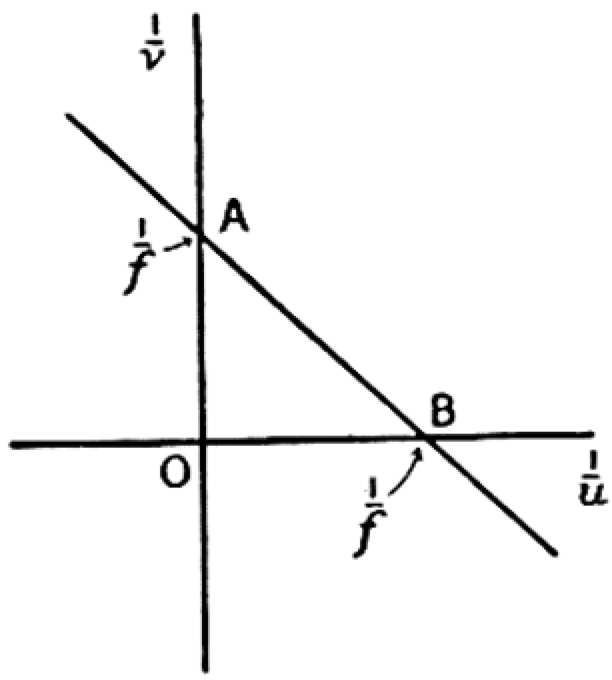


FIG. 541

Method 4. *Magnification.*—The magnification of a real image may be measured quite simply. The diameter of the gauze-covered hole, or similar object, is first measured. This is OO′. Then the diameter of the image of this hole formed on the screen is measured either with the help of dividers or by using a screen covered with squared paper. This is II′. The magnification is then numerically equal to $\frac{II'}{OO'}$. This number is found for various distances of the object and image, and so the corresponding values of m to which it is numerically equal can be written down. Since we are confining the observations to real objects and images, m is always a negative number. We have now

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$$

and

$$\frac{v}{u} = -m$$

so that

$$\frac{1}{v} - \frac{m}{v} = \frac{1}{f}$$

or

$$v = (-m)f + f$$

Therefore if a series of readings of II′ and v is taken, we can obtain a straight line graph (Fig. 542) by plotting v against $(-m)$, which is always positive in the observed region. The slope of the graph is equal to f, and its intercept on the $(-m)$ axis is -1. The intercept on the axis of v is equal to f. Note that $(-m)$ is equal to $+1$ when v is equal to $2f$, *i.e.* the object and image are both at the centre of curvature.

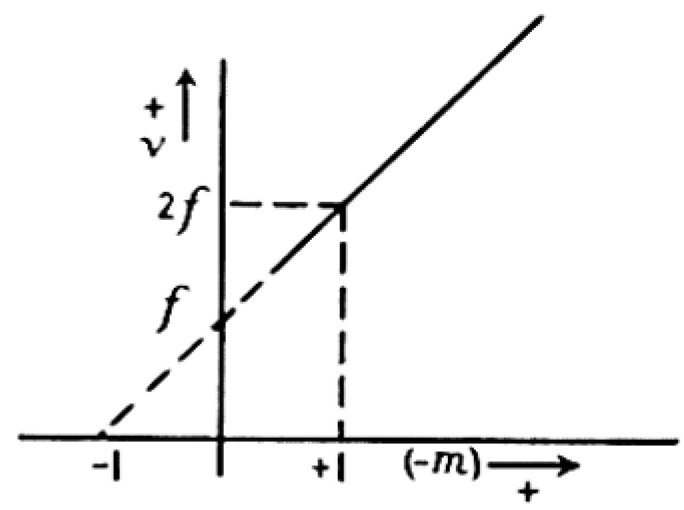


FIG. 542

It is also possible to deduce f from a graph of u against $\dfrac{1}{(-m)}$.

Determination of f and r for a Convex Mirror.—In general these determinations are less direct and accurate than those for a concave mirror because real objects do not give rise to real images.

Method 1. *Distant Object.*—The image of a distant object situated behind the observer's head is viewed in the mirror and located by eliminating parallax between it and a pin viewed over the mirror. The distance from the pin to the mirror is f.

Method 2. *Measurement of u and v.*—A series of simultaneous values of u and v is taken, the image being located by the method of no parallax in each case. The readings may then be used in any of the three ways described under (*a*), (*b*) and (*c*) for a concave mirror on pages 731 and 732.

In an experiment of this kind the position of the virtual image may be found with the help of a plane mirror rather more easily than by locating it directly. The arrangement is shown in Fig. 543. The plane mirror M_2 is placed in front of the convex mirror M_1 so as to cover, say, the lower half of the latter. Thus two virtual images of the pin O will be formed, one in each mirror, and by adjusting the relative positions of O and the two mirrors it is possible to make these images coincide in space at I as judged by the absence of parallax when they are viewed by an eye placed as shown. The position of the images is then easily obtained, because we know that for the plane mirror IM_2 is equal to OM_2. The latter distance can be measured, and we have, for M_1,

FIG. 543

$$IP = IM_2 - M_2P$$

$$= OM_2 - M_2P$$

while the object distance OP can be measured directly. On substituting in the mirror equation u must be put equal to $+OP$ and v to $-IP$.

Method 3. *Virtual Object.*—(*a*) A convex lens (Fig. 544) is arranged so as to form an image of a luminous object O on a screen S at I. The convex mirror M is then inserted between L and S so as to face L, and it is moved to and from L until the image of O is formed beside O itself. When this state of affairs is attained, the light reaching M from the lens is converging towards the centre of curvature of M because it is reflected back along its own path and the rays must therefore be normal to the reflecting surface. It follows, therefore, that PI is equal to the radius of curvature, which can be determined by subtracting LP from LI. It is necessary that LS should be greater than the radius, and this entails placing O not too far from L.

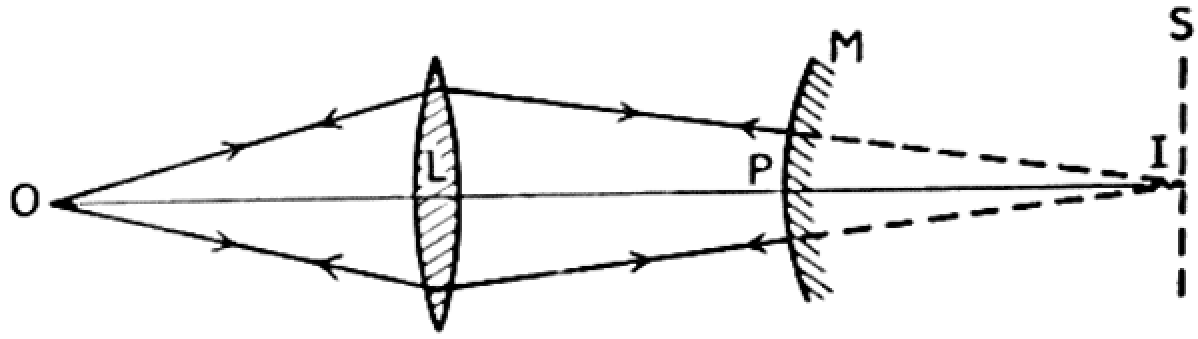


FIG. 544

(*b*) A convex lens is made to form a real image I_1 of an object O (Fig. 545). This is located by a screen S_1, and LI_1 is measured. Then the mirror M is interposed so as to reflect the light before it arrives at I_1 and forms a real image I_2, which can be received on a screen placed, say, at S_2. The distance I_2P is then measured. We then have, for the mirror,

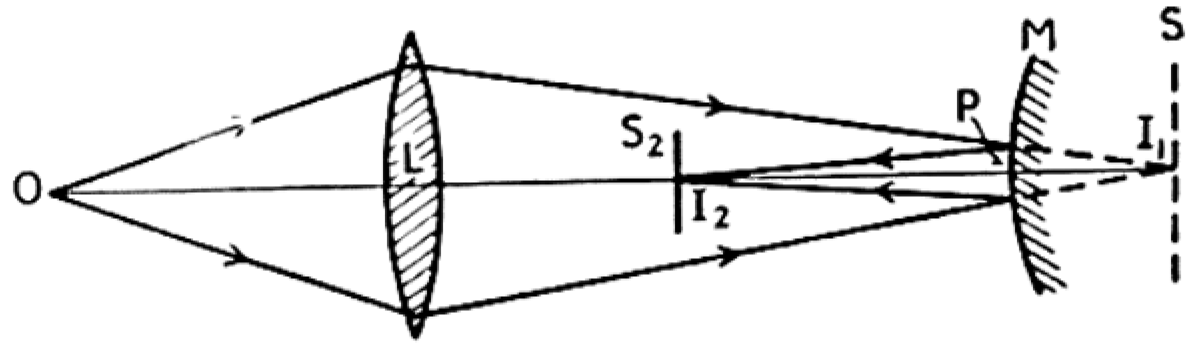


FIG. 545

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

where

$u = -I_1P = -(LI_1 - LP)$ the object being virtual

and

$v = +S_2P$ the image being real

The final real image will not be formed unless the virtual object I_1 is within the focal length of the mirror.

(*c*) The previous experiment may be carried out using a concave mirror instead of a convex lens to form the virtual object. The arrangement is shown in Fig. 546 and the calculation is the same as in (*b*).

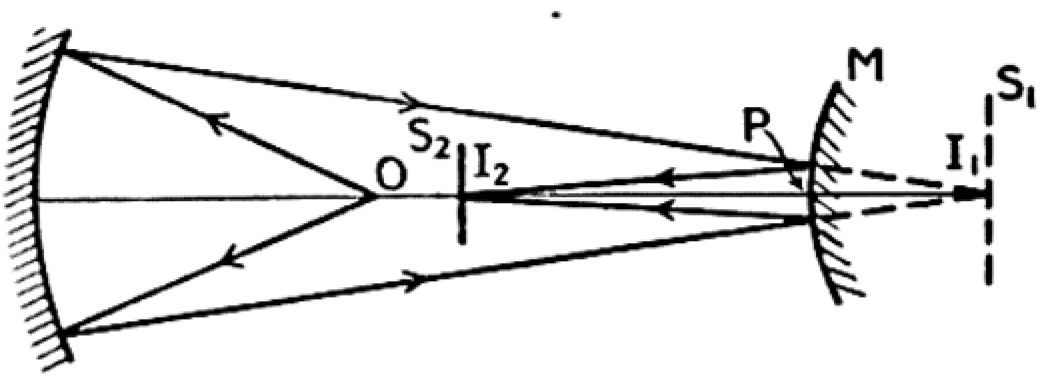


FIG. 546

5. DEPARTURE FROM PARAXIAL CONDITIONS

It must be emphasized once again that all the foregoing theory and experiments are applicable only when the width of the mirror is small compared with its radius of curvature and with the distances of the object and image, so that the rays are close to the axis. We now mention what happens in two cases in which this condition is not fulfilled.

Wide Concave Spherical Mirror.—In Fig. 547 the mirror extends sideways from its principal axis much further in comparison with its radius of curvature than in the cases we have considered previously. The reflected rays formed by incident rays parallel to the axis may easily be traced using the laws of reflection, and when this is done we see immediately that, as the incident rays get further from the axis, the point at which their reflected rays cross the axis moves from the principal focus F towards the pole of the mirror.

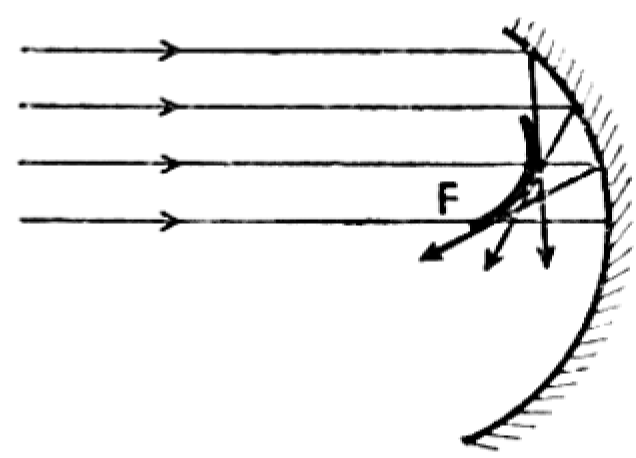


FIG. 547

Thus no definite point image is formed. Curved lines can be drawn from F such that each reflected ray in turn is tangential to them. They constitute what is called a **caustic curve** along which there is a concentration of light because the intersections of adjacent reflected rays lie on it. Of course with a spherical mirror the caustic is the *surface* generated by the curve when Fig. 547 is rotated about the principal axis. A caustic curve can often be seen on the surface of the liquid in a teacup. It is formed by reflection at the curved surface of the cup above the tea. The lack of clarity in the image which occurs when rays too distant from the pole of the mirror are included in its formation is called **aberration.** It is reduced by placing a screen with a hole in it (a **diaphragm**) in front of the mirror so as to exclude all but paraxial rays. Evidently, if we wish to concentrate all the reflected rays at one point, the curvature of the mirror should diminish towards its rim. It can be shown that for parallel incident rays a parabolic instead of a circular curve fulfils this condition to exactly the right extent for perfect focusing (Fig. 548). Parabolic mirrors are used as searchlight reflectors with the source of light at the focus. It should be mentioned that the parabolic mirror gives perfect image formation only when the rays are as in Fig. 548. For

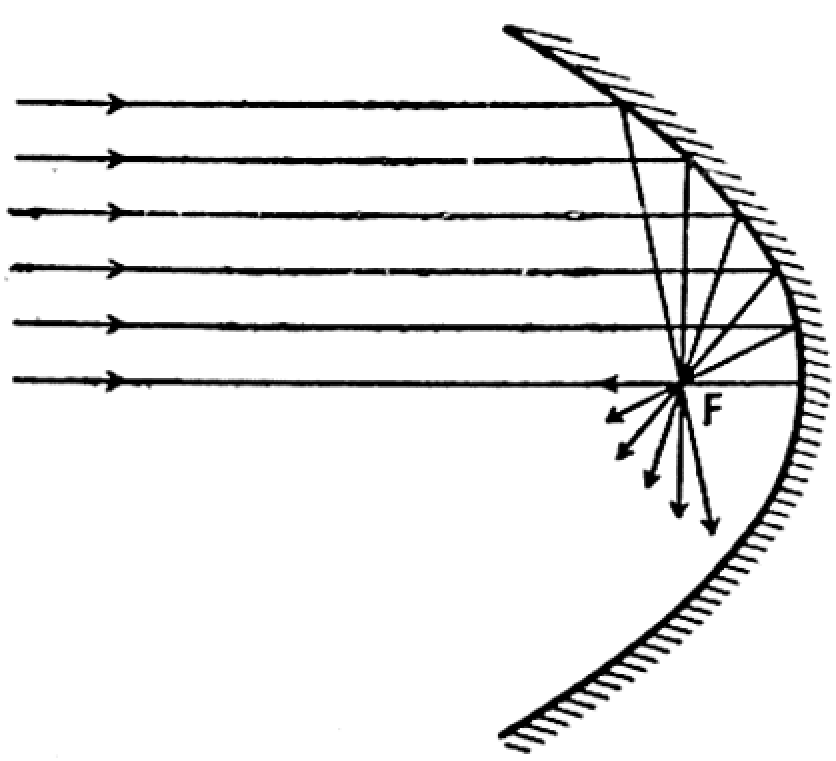


FIG. 548

positions of the object other than at infinity or the principal focus there is still aberration unless an **ellipsoidal** mirror is used, the conjugate points being at its two foci.

Image of an Object considerably off the Axis of a Concave Mirror. Focal Lines.—Fig. 549 is a perspective drawing of the reflection of four rays incident on a concave mirror from a point O which is a considerable distance from the principal axis CP. The reflected beam does not come to a focus at any one point but forms two **focal lines** HI and JK. The beam is said to be **astigmatic.** We cannot discuss the matter fully and quantitatively, but a clue to the formation of the astigmatic pencil lies in the lack of symmetry of the axis of the incident pencil with respect to the mirror. We have previously dealt with objects which are so near to the axis that this lack of symmetry is negligible.

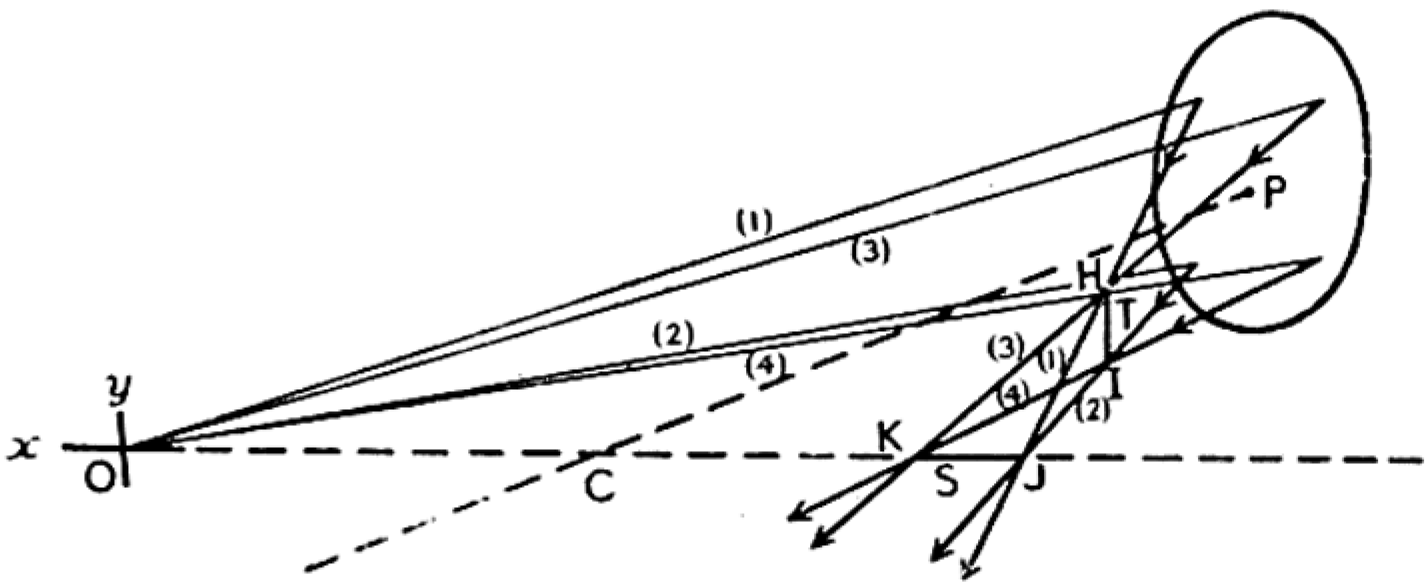


FIG. 549

The normal to the mirror at any point of incidence is a line passing through C. Therefore, for any incident ray the plane of incidence (*i.e.* the plane containing the incident ray, the normal and the reflected ray) must contain the points O and C. Thus all the planes of incidence intersect in the axis OC, and it follows that all the reflected rays cross OC (produced).

Suppose that the incident rays (1) and (2) are equally inclined to OC. From this it follows that their angles of incidence are equal and so therefore are their angles of reflection. Consequently, the reflected rays (1) and (2) cross the axis OC at the same point J. Rays (3) and (4) are a similar pair which cross OC together at K, but K and J do not coincide as they would (approximately) if O were very near to the principal axis. There is a series of such points along OC, each one being the point of intersection of all those rays whose angles of incidence are the same. Rays (1) and (3) have a common *plane* of incidence and their reflected rays cross each other at H. Similarly, the reflected rays (2) and (4) which also have a common plane of incidence cross at another point I. Therefore another focal line, HI, is also present which has the property that any one point on it is the point of intersection of reflected rays whose *planes* of incidence are the same (as distinct from common *angles* of incidence for points on the line JK).

It will be noticed that a given pair of reflected rays do not intersect at all unless they have either a common plane or angle of incidence, in which case they intersect on one or other of the focal lines.

Thus if a very small luminous point is situated at O, it is possible to obtain two line images on a screen which is placed in turn in the planes of the focal lines.

If x (Fig. 549) is a line object lying along OC it can be regarded as a series of points, each of which produces its own pair of focal lines. This will have the effect of lengthening the focal line JK and producing a series of focal lines parallel and side by side with HI. Thus HI is no longer a single line but a rectangle, and is therefore not a definite image at all.

Similarly, a line object, y, perpendicular to the first would be focused at HI but not at JK. Therefore an illuminated piece of wire gauze or a pair of crosswires gives an image at each focal line, its vertical wires (say) being in focus at one position and the horizontal ones at the other.

The properties of an astigmatic beam are more fully discussed on pages 819–821, where the corresponding effect in lenses is dealt with.

EXAMPLES XLIII

1. Find the relation connecting the focal length of a convex spherical mirror with the distances from the mirror of a small object and the image formed by the mirror.

A convex mirror, radius of curvature 30 cm., forms a real image 20 cm. from its surface. Explain how this is possible, and find whether the image is erect or inverted. (L.I.)

2. Deduce a formula connecting u, v and r, the distances of object, image and centre of curvature from a spherical mirror.

A mirror forms an erect image 30 cm. from the object and twice its height. Where must the mirror be situated? What is its radius of curvature? Assuming the object to be real, determine whether the mirror is convex or concave. (L.H.S.)

3. Define focal length of a spherical mirror, and describe how to find the focal length of a convex mirror.

A concave mirror forms a real image 3 times as long as the object; when the latter is moved 4 in. the image is only 1·5 times as long. What is the focal length of the mirror? (L.I.)

4. A convex mirror forms (*a*) a virtual image half the length of the object, (*b*) a real image twice the length of the object, which is, in each case, perpendicular to the axis of the mirror. If the focal length of the mirror is 15 cm., find the positions of the object and image in each case, and draw ray diagrams showing the paths of a pencil of rays from a non-axial point on the object to the eye of an observer. (L.I.)

5. What is the nature of the image formed by a convex mirror when an object is placed on its axis?

A convex mirror produces an image whose height is one-third that of the object, the distance between the object and image being 32 cm. Calculate the radius of curvature of the mirror, and indicate what sign convention you adopt in the calculation. (L.Med.)

6. Rays parallel to the principal axis of a concave mirror are incident on the mirror at various distances from the axis. Draw a diagram showing the paths of the rays after reflection. Define the principal focus of the mirror and deduce its position.

A linear object 1 in. long is placed along the axis of a concave spherical mirror whose focal length is 10 in. If the end of the object further from the mirror is 40 in. from its surface, what is the length of the image? (L.I.)

7. Explain, with the aid of ray diagrams, the formation of real and virtual images by reflection from a concave spherical mirror.

Find the focal length of a spherical mirror such that whether an object is placed at a distance of either 8 or 16 cm. away from the mirror, an image is formed of the same size in each case. Find also the magnification of the image. Is the mirror convex or concave? (L.Med.)

8. A small object on the axis of a convex mirror, at distance u from the pole, gives an image at distance v from the pole; obtain a formula connecting u, v, and the radius of curvature of the mirror.

Describe how you would measure the radius of curvature of a convex mirror experimentally.

A parallel beam of rays, of diameter 5 cm., falls on a convex mirror of radius of curvature 30 cm., whence it is reflected on to a wall 2 m. from the mirror, outlining a circular patch of light on the wall. Find the diameter of this patch of light. (O.H.S.)

9. Deduce from first principles the relation between the object distance, the image distance, and the focal length of a spherical mirror. Draw diagrams showing the changes in size and position of the image as an object is moved along the axis of a concave mirror up to its surface from a long distance away.

How would you determine experimentally the focal length of a concave spherical mirror?

Two illuminated objects are placed 15 cm. apart in a line perpendicular to the visual axis of, and 60 cm. in front of, an eye. Calculate the position and the distance between the images reflected from the surface of the cornea, which has a radius of curvature of 8 mm. (L.A.)

10. A concave mirror of 15 ft. focal length reflects the light of the moon to a smaller convex mirror facing it, 12 ft. in front, which in turn reflects it to form a real image in the middle of the big mirror itself. Calculate the focal length of the convex mirror and the number of diameters it magnifies what would have been the original real image. (L.Med.)

11. Prove the relation connecting the distances of object and image from a concave spherical mirror with the radius of curvature of the mirror.

A concave mirror forms a real image 1·25 cm. long of an object 1 cm. long placed at right angles to its axis. When the object is moved a little towards the mirror, which itself remains fixed, the screen has to be moved 18·0 cm. to refocus the image which is now 2·15 cm. long. What is the focal length of the mirror? (L.I.)

12. A convex mirror, a plane mirror, and a pin are placed in order on an optical bench so that the two images of the pin formed by the two mirrors coincide. If the pin is 20 cm. from one mirror and 30 cm. from the other, find the focal length of the convex mirror. (L.I.)

13. Define the *focal length* of a spherical mirror, and show how it is related to the radius of curvature of the mirror.

A spherical mirror of focal length 15 cm. is set up so as to intercept rays of light which are converging towards a point on its axis 20 cm. behind its silvered surface. Where will an image be formed if the mirror is (*a*) concave, (*b*) convex? State the sign convention which you use. (L.I.)

14. A concave spherical mirror is circular in shape, measuring 3 cm. across, and has a radius of curvature of 100 cm. A screen is placed at its centre of curvature perpendicular to its principal axis, and a point source of light is interposed on the axis between the mirror and the screen. Calculate the diameter of the patch of light formed on the screen by reflection, if the source is (*a*) midway between the screen and the mirror, and (*b*) 75 cm. from the mirror. (L.I.)

15. Explain the formation of the "caustic curve" formed when parallel light is reflected by a concave mirror of large aperture, and define the principal focus of a concave mirror.

A ray of light parallel to the axis of a concave mirror of 6 in. radius of curvature is incident at a point on its surface whose perpendicular distance from the axis is 3 in. After reflection the ray crosses the axis at a point I. Find the distance between I and the principal focus of the mirror. (L.I.)

Chapter XLIV

REFRACTION

1. INTRODUCTION

The Laws of Refraction.—It has already been stated that when a ray of light passes from one medium to another a portion of the light energy is reflected back into the first medium while the remainder passes into the second with its direction changed. We now have to study this deviation of the transmitted ray which is known as refraction.

In Fig. 550, AB is a ray travelling in medium (1) and striking the boundary of medium (2) at B. The reflected ray is BD and the refracted ray is BC. If N_1BN_2 is a normal to the boundary at B, the angle of incidence is $A\widehat{B}N_1$ $(=i_1$, say) and the angle of refraction is $C\widehat{B}N_2$ $(=i_2$, say).

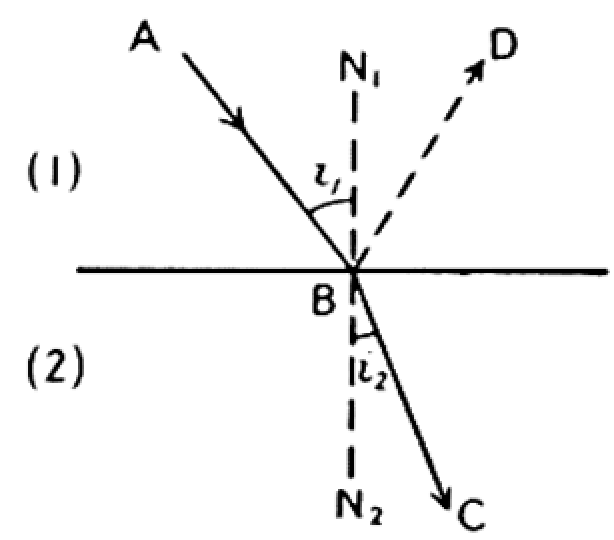


FIG. 550

In the diagram i_2 is shown as being smaller than i_1, in which case the light is said to have been refracted "towards the normal." The other case of bending "away from the normal" is equally possible, according to the relative properties of the two media. The second medium is said to be optically denser when the bending is towards the normal and optically rarer in the other case. It is an important experimental fact that if an incident ray is directed along CB in medium (2), it will pass into medium (1) along the path of BA. Thus the path of the light is reversible.

The laws of refraction, which correlate all experimental observations, can be stated as follows:—

1. The incident ray, the refracted ray, and the normal at the point of incidence all lie in the same plane.

2. For a given pair of media, and for light of a given wave-length, the ratio of the sine of the angle of incidence to the sine of the angle of refraction is a constant independent of the angle of incidence. This is known as **Snell's law.**

The first law is identical with the first law of reflection. The second law can be written mathematically as

$$\frac{\sin i_1}{\sin i_2} = {}_1n_2$$

where ${}_1n_2$ is a constant for light going from medium (1) to medium (2) and is called the **refractive index** from medium (1) to medium (2).

Refractive index is a very important physical quantity and evidently depends upon the nature of both the media concerned. It is therefore much more satisfactory to replace the number ${}_1n_2$ by a combination of two numbers, one of which is specific to the first medium of the pair and the other to the second. This can be done by defining the **absolute refractive index** of a medium as the refractive index when light passes from a vacuum into the medium. Thus in Fig. 550, if medium (1) is a vacuum, the absolute refractive index of medium (2) (n_2) is given by

$$\frac{\sin i_1}{\sin i_2} = n_2$$

It is evident that by doing this we are making the refractive index of the vacuum itself equal to unity, because if medium (2) is a vacuum as well as medium (1), there is no deviation of the incident ray and i_1 is equal to i_2.

In practice it is often sufficiently accurate to use air instead of a vacuum, the refractive index of air being about 1·0003.

If we determine the absolute refractive index of a certain medium (1) and find it to be, say, n_1 and that of a second medium (2) to be n_2, and we then make a determination of the refractive index (${}_1n_2$) for light passing from (1) to (2), it is invariably found that

$${}_1n_2 = \frac{n_2}{n_1}$$

Therefore we can always calculate the refractive index for a pair of substances if we know their absolute refractive indices against a vacuum. From now on we shall refer to the absolute refractive index of a substance simply as its refractive index.

Referring to Fig. 550, we can now write for any pair of media

$$\frac{\sin i_1}{\sin i_2} = \frac{n_2}{n_1}$$

or

$$n_1 \sin i_1 = n_2 \sin i_2 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

which is an excellent way of remembering the relationship. If either of the two media is a vacuum, its refractive index is unity.

The above relationship is independent of the direction of the light owing to the experimental fact that if the direction of the light is reversed, the angles are unchanged.

We next consider the passage of a ray of light from a medium of refractive index n_1 into and through one or more parallel-sided slabs of different materials (n_2, n_3, etc.) and thence into the first medium again (Fig. 551).

For the first refraction,

$$n_1 \sin i_1 = n_2 \sin i_2$$

The two angles marked i_2 are equal because the normals are parallel to each other. Therefore for the second refraction,

$$n_2 \sin i_2 = n_3 \sin i_3$$

Similarly for the third refraction,

$$n_3 \sin i_3 = n_1 \sin i_4$$

When these three equations are added together, we obtain the equation

$$n_1 \sin i_1 = n_1 \sin i_4$$

which gives

$$i_1 = i_4$$

This proves that the initial and final directions of the ray are parallel to each other no matter how many slabs of material the light has passed through. There is however a sideways displacement. The same result applies, of course, when there is only one plate of material with the same medium on both sides of it, *e.g.* a plate of glass in air or water.

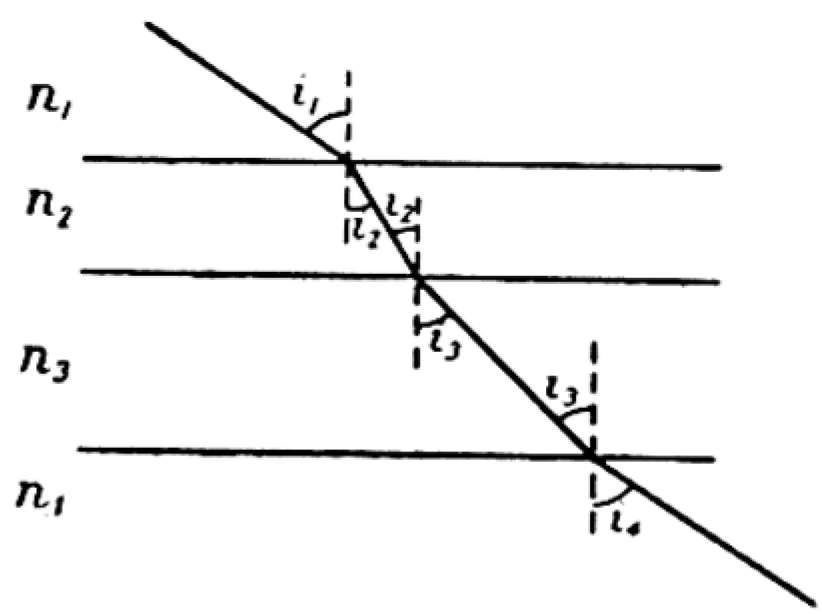


FIG. 551

The student should refer to Section 2 of Chapter XLI for a discussion of refraction in terms of the wave nature of light. It is shown on page 696 that the relationship between the angles i_1 and i_2 is

$$\frac{\sin i_1}{\sin i_2} = \frac{c_1}{c_2}$$

where c_1 and c_2 are the speeds of propagation of the light in medium (1) and medium (2) respectively.

Comparing this with equation (1) gives the relation

$$\frac{c_1}{c_2} = \frac{n_2}{n_1}$$

which means that the speed of propagation of light in a medium is inversely proportional to the refractive index of the material.

It should be emphasized that refractive index (and therefore speed of propagation) depends upon the wave-length (colour) of the light.

2. REFRACTION BY PLANE SURFACES

Formation of an Image by Refraction at a Plane Surface.—In Fig. 552, O is a point object in a medium of refractive index n_1 which is separated from another medium n_2 by a plane boundary. It is assumed

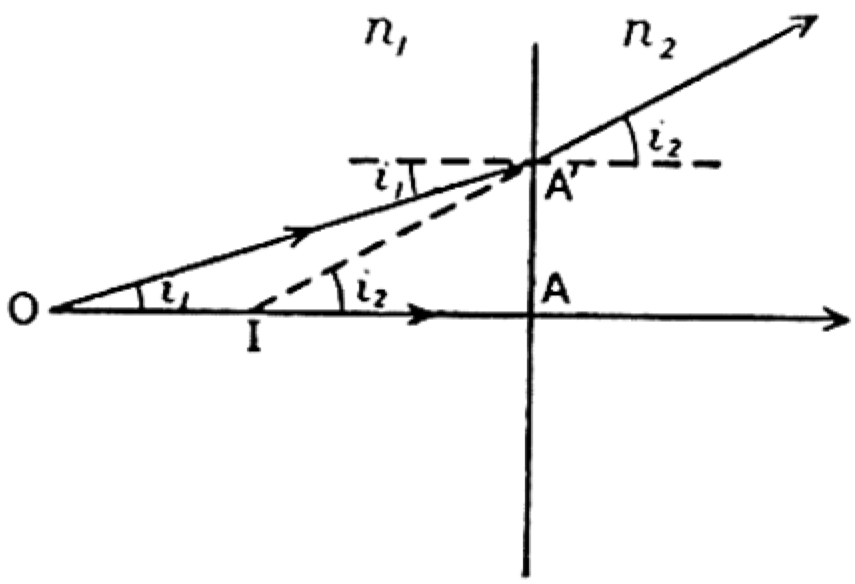


FIG. 552

that $n_2 < n_1$, so that light entering the second medium is bent away from the normal. The ray OA which strikes the boundary normally (angle of incidence zero) passes into the second medium undeviated, but OA′ is deviated according to the equation

$$n_1 \sin i_1 = n_2 \sin i_2$$

Let the refracted ray be produced backwards to meet OA at I. Then since the normal at A′ is parallel to OA we have

$$A'\widehat{O}A = i_1$$

and

$$A'\widehat{I}A = i_2$$

$$\therefore \sin i_1 = \frac{AA'}{OA'}$$

and

$$\sin i_2 = \frac{AA'}{IA'}$$

Substituting in the original equation gives

$$n_1 . \frac{AA'}{OA'} = n_2 . \frac{AA'}{IA'}$$

or

$$\frac{OA'}{IA'} = \frac{n_1}{n_2}$$

This means that in general I is *not* a fixed point in OA for every possible incident ray. If, however, we confine ourselves to rays which are only

slightly inclined to OA (*i.e.* rays for which AA′ is small compared with OA and IA), we can write

$$OA' = OA \quad \text{approximately}$$

and

$$IA' = IA \quad \text{approximately}$$

so that

$$\frac{OA}{IA} = \frac{n_1}{n_2} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

which means that for all incident rays which satisfy the above condition the refracted rays diverge from the same point I.

Now the pupil of the eye is sufficiently small to ensure that an eye placed on OA produced (Fig. 553) will receive only a narrow pencil of rays, for each of which AA′ is small and the last equation holds good. An image of O is then seen at I whose position is given by this equation. It must be realized that the width of the pencils is greatly exaggerated in the drawing.

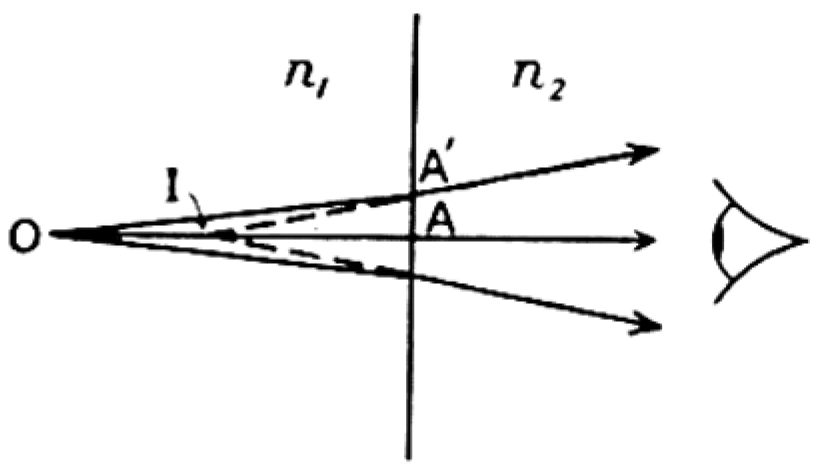


FIG. 553

The distance IA is called the **apparent depth** of O within the first medium when it is viewed from the second medium by an eye situated on the normal through O.

If O is situated in glass ($n_1 = 1{\cdot}5$, say) and the eye is in air ($n_2 = 1$) we have

$$IA = \frac{n_2}{n_1} \times (OA) = \tfrac{2}{3}OA$$

From this we conclude that the apparent thickness of a glass plate viewed normally is two-thirds of its real thickness, while the apparent depth of a quantity of water (again when viewed normally) is three-quarters of its real depth ($n_1 = 1{\cdot}33$ or $\frac{4}{3}$, $n_2 = 1$).

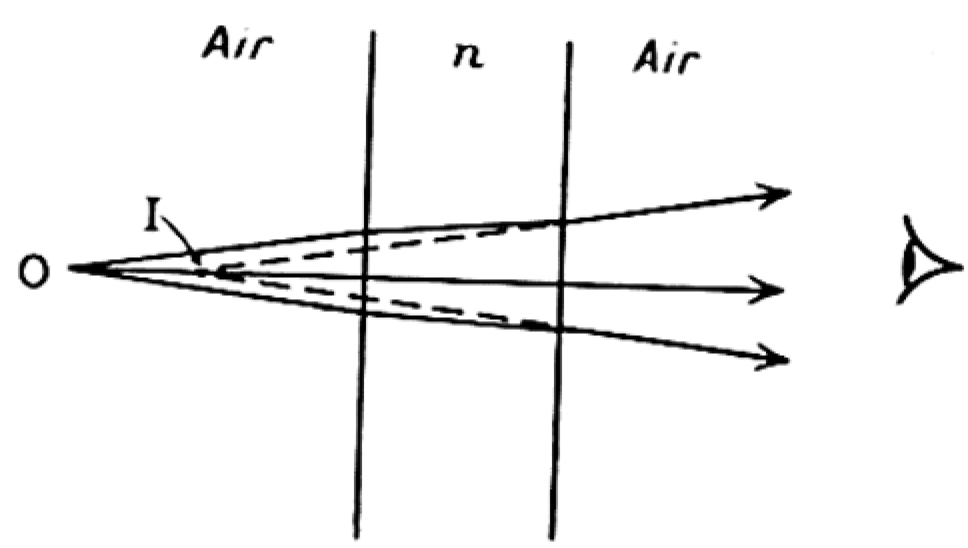


FIG. 554

The student should be able to prove for himself, by drawing a ray diagram, that when an object in air is viewed through a parallel-sided slab of material of refractive index n and thickness t (the faces of the slab being perpendicular to the line joining the object to the eye), it appears to be nearer to the eye by an amount equal to the difference between the real and apparent

thickness of the slab, *i.e.* by $\left(t - \frac{t}{n}\right)$ (Fig. 554). This is true regardless of the position of the slab between the object and the eye.

Example.—*A vertical sheet of glass 3 cm. thick separates air from water. An eye situated in the air looks normally through the glass at a small object situated in the water 24 cm. from the glass/water surface. Find the apparent position of the object. (Refractive index of glass* $= \frac{3}{2}$, *of water* $= \frac{4}{3}$.)

The ray diagram is shown in Fig. 555. A narrow pencil of rays travels from the object O, and at the glass/water surface the rays are bent towards the normal because they are entering a medium of higher refractive index. Therefore the

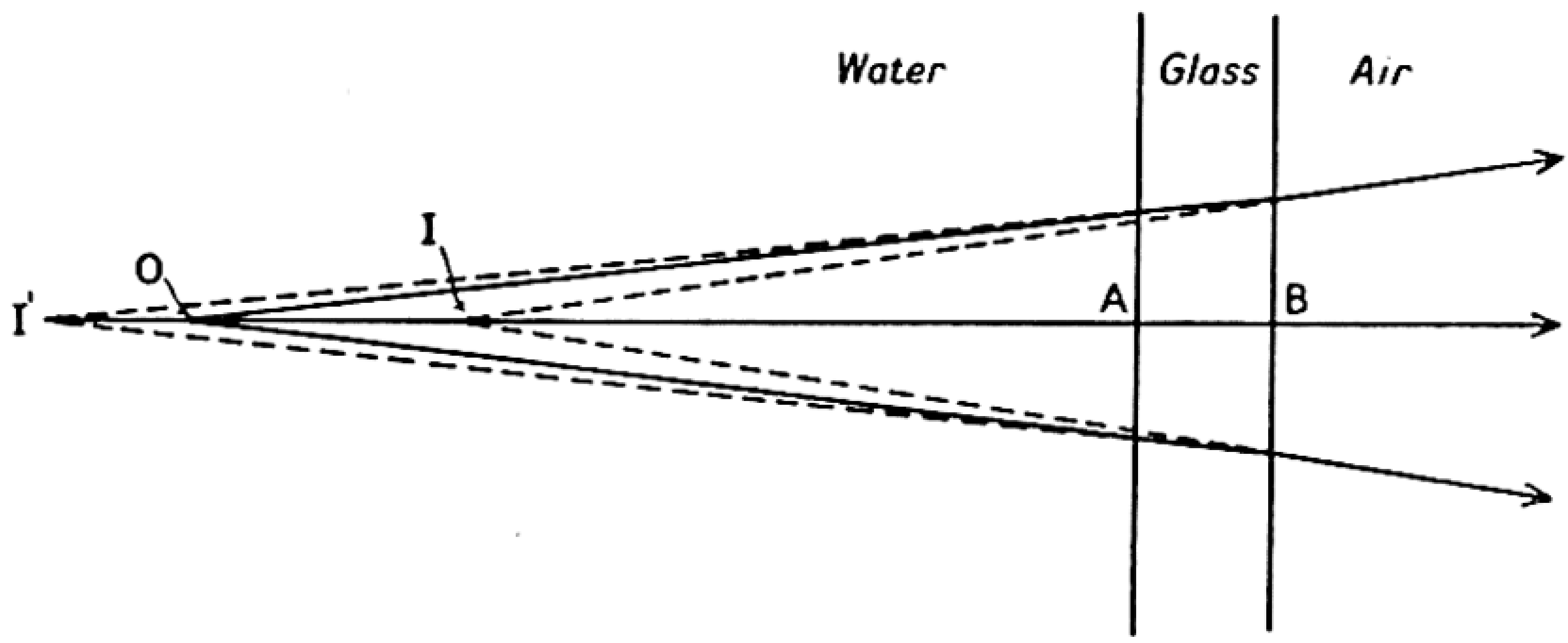


FIG. 555

point I′ from which the rays appear to diverge after the first refraction is further from A than O is, and its position is given by equation (2) on page 744, *i.e.*

$$\frac{\text{OA}}{\text{I'A}} = \frac{\frac{4}{3}}{\frac{3}{2}} = \frac{8}{9}$$

so that

$$\begin{aligned} \text{I'A} &= \frac{9}{8}.\text{OA} \\ &= \frac{9}{8} \times 24 \\ &= 27 \text{ cm.} \end{aligned}$$

The rays in the glass are therefore diverging from the point I′, which is 30 cm. from the glass/air surface where the second refraction occurs. The image I′ acts as an object for this refraction and, by the same equation, the position of the final image I is given by

$$\frac{\text{I'B}}{\text{IB}} = \frac{\frac{3}{2}}{1} = \frac{3}{2}$$

$$\begin{aligned} \text{IB} &= \frac{2}{3} \times \text{I'B} \\ &= \frac{2}{3} \times 30 \\ &= 20 \text{ cm.} \end{aligned}$$

The final image, *i.e.* the apparent position of the object, is therefore 20 cm. from the glass/air surface, or 17 cm. from the glass/water surface.

Caustic Curve Due to Refraction.—We now consider the refraction of *all* the rays striking a plane surface from a point object. By choosing various angles of incidence for light rays diverging from O (Fig. 556) and striking the plane boundary, we can easily calculate the directions of the refracted rays when n_1 and n_2 are given. The result is as shown in the figure, in which it is again supposed that n_1 is greater than n_2 (*e.g.* when the first medium is glass or water and the second is air). It will be noticed at once that the refracted rays produced backwards cross the normal OA at points which approach the refracting surface as the angle of incidence increases. The position which O appears to occupy to an eye placed in the second medium depends upon which of the refracted rays are being received, *i.e.* upon the position of the eye. Thus an eye receiving rays (2) and (3) would see an image of O at their intersection (when produced backwards), and the point is different for each pair of adjacent rays. The thick curved line is a **caustic curve** which is drawn so that each dotted line is tangential to it. There is, of course, a similar curve on the other side of OA, and the point (or **cusp**), where the two sections of the caustic meet, is the position of the image for light passing normally through the boundary. If we regard the image of O as being the point of intersection of two *adjacent* rays entering the eye (as it would be if the pupil of the eye were very small), then we see that the caustic curve is the locus of the image, because the nearer the two refracted rays approach each other the nearer does their point of intersection approach the curve.

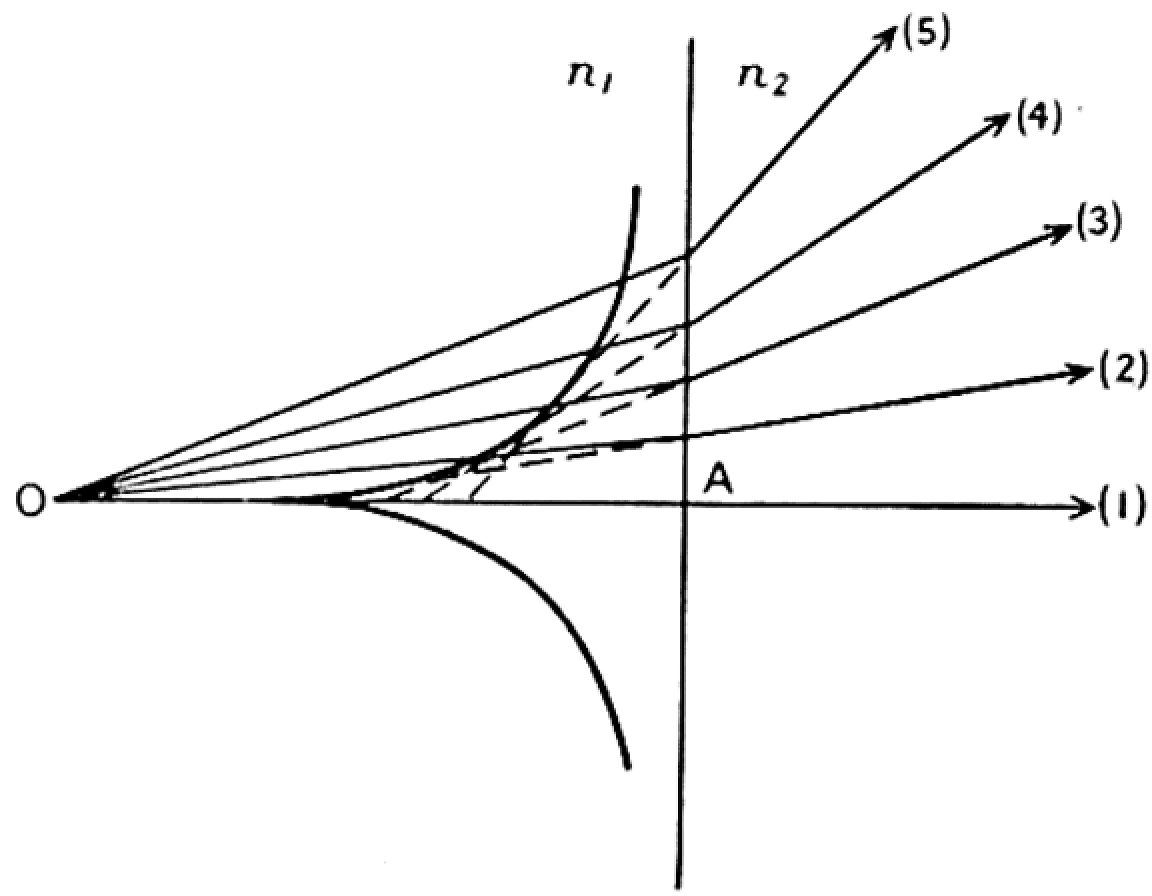


FIG. 556

The caustic curve can be constructed experimentally by means of a glass block and pins. The block is laid flat on a piece of drawing-paper, its outline marked, and a pin P_0 (Fig. 557) is stuck in the paper right against one of its faces. This pin is viewed through the other face of the block, and two pins such as P_1 and P_2 are placed so as to be in line with its image. The plotting of refracted rays in this way is repeated until a series is obtained. After the block is removed, the rays are produced

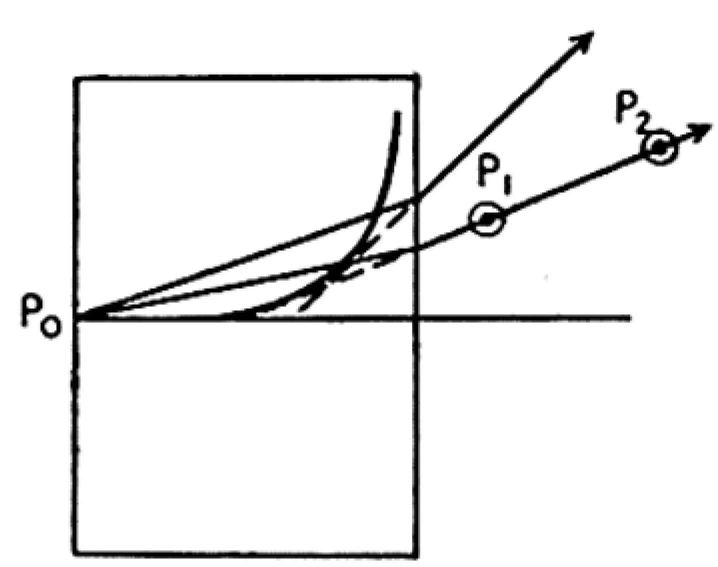


FIG. 557

backwards and the caustic curve carefully drawn in such a way that the lines are tangential to it.

The change of position of the image with the position of the eye can be readily observed when a bright object such as a coin is placed on the bottom of a bowl filled with water and viewed by an eye, which is slowly moved from a position vertically above the object. The image moves along the caustic curve and the depth of the bowl appears to diminish as the eye moves from the normal position (Fig. 558). For the same reason a tank of water of uniform depth appears to be shallower at the end further from the observer.

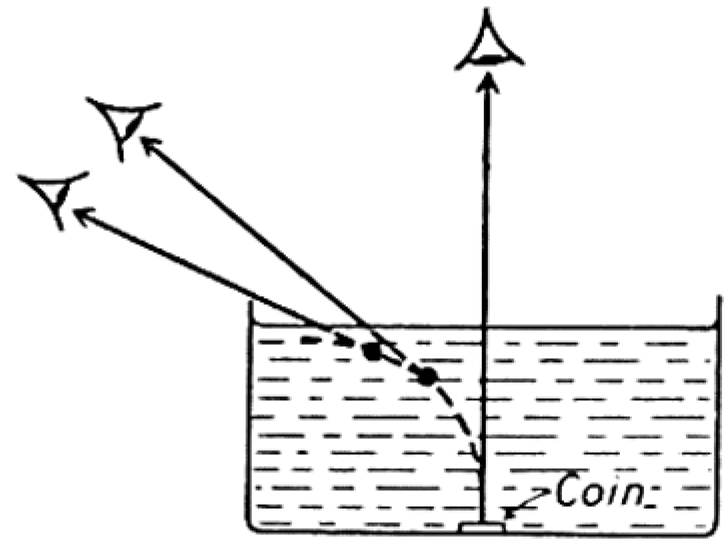


FIG. 558

Atmospheric Refraction.—The refractive index of a gas increases when its density is increased by either increased pressure or decreased temperature. The relationship can be represented by the equation

$$\frac{n-1}{\text{density}} = \text{constant}$$

Evidently, therefore, the refractive index of the atmosphere under normal conditions diminishes with height above the earth's surface. The effect of this on light coming from outside the atmosphere is best understood by a consideration of its wave nature. The analogous case for sound is described on page 600 (Vol. 3).

Suppose that a plane light wave arrives at the earth's atmosphere from a star or the sun. In Fig. 559, A_1B_1, which is at right angles to the parallel rays, represents a plane wave-front just before the light reaches the earth's atmosphere. After passing through a certain amount of atmosphere the wave-front has moved forward and has been rotated slightly to the position A_2B_2. The reason for the rotation is that, on account of the lower altitude of the path B_1B_2, the refractive index here is higher and the speed of propagation is therefore lower than along A_1A_2. The wave-front is therefore deviated towards the denser region, which means that the waves or rays travel along the curved paths shown in the figure. An observer situated at E would judge the light to be coming along the dotted straight path and, if he made a measurement, would obtain too high a value for the elevation of the star. The deviation will obviously be zero for a star at its zenith, and a maximum for a star or the sun setting on the horizon.

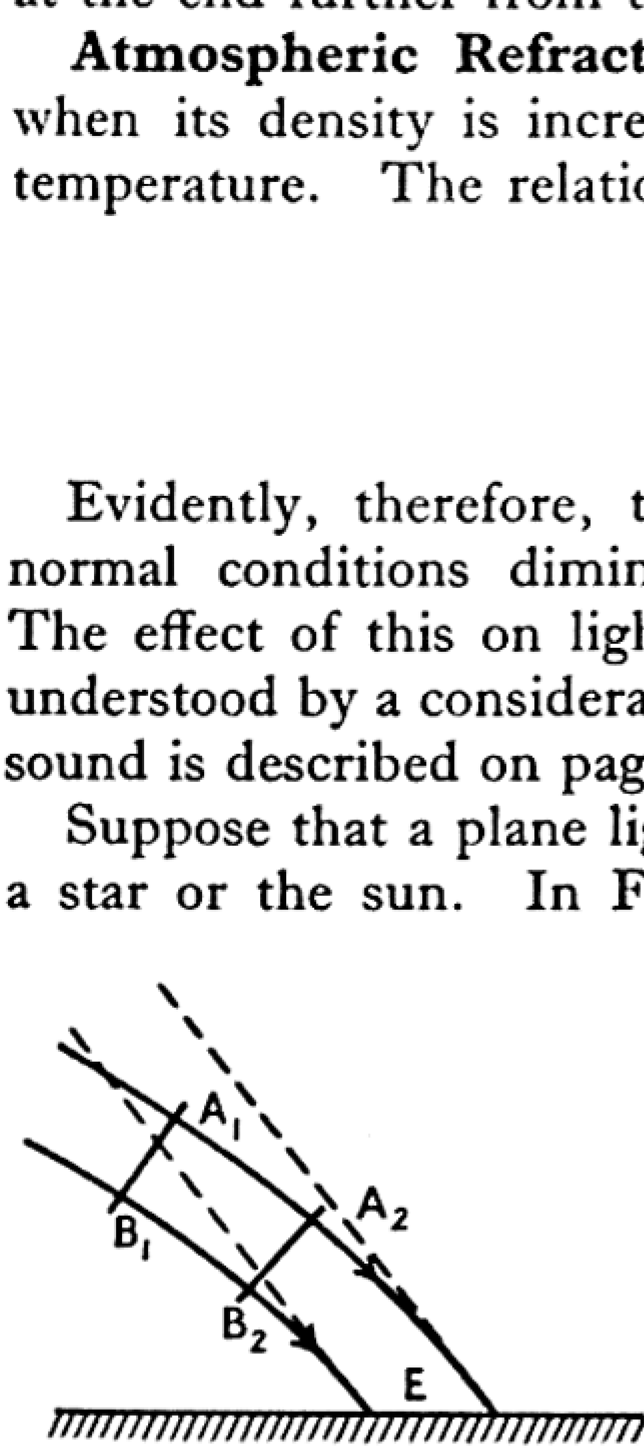


FIG. 559

A smaller-scale phenomenon due to the variation of refractive index with temperature is to be seen on tarred roads in hot weather. The temperature of the road surface is higher than the average temperature of the atmosphere, so that the lower layers of air are warmer than those above. This produces an upward bending (Fig. 560) of light which is travelling approximately horizontally. The light received at E therefore appears to have been reflected off the road itself and gives the impression that the road is covered with water. The phenomenon is particularly noticeable when the observer is ascending a slope to a horizontal portion of the road, because his eye is then on a level with the horizontal surface. Similar effects on a bigger scale in a sandy desert give the impression that objects such as trees, which are actually situated at A, are at A′. The phenomenon is known as a **mirage.**

FIG. 560

3. TOTAL REFLECTION

Critical Angle.—We now discuss a phenomenon which occurs when light passes from a medium of refractive index n_1 into another of *lower* refractive index n_2. The law of refraction states that

$$n_1 \sin i_1 = n_2 \sin i_2$$

so that when n_2 is less than n_1, i_2 must always be greater than i_1, and the light is bent away from the normal (Fig. 561). Some of the light is reflected back into the first medium as indicated by the dotted ray.

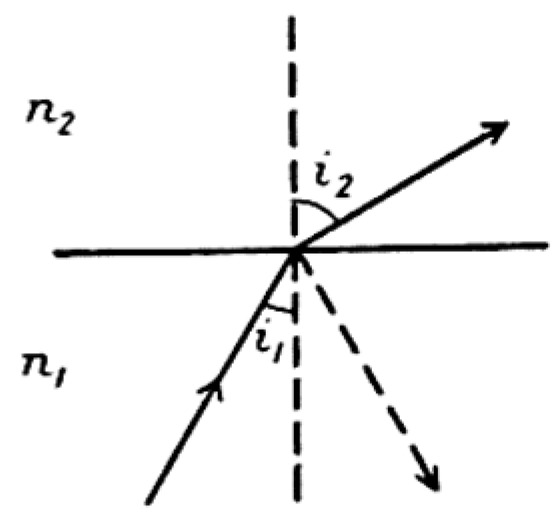


FIG. 561

Thus, when the angle of incidence (i_1) is gradually increased, the angle of refraction (i_2) eventually reaches 90° for a value of i_1 (i_c, say), which is less than 90°, and is given by

$$n_1 \sin i_c = n_2 \sin 90°$$

that is

$$\sin i_c = \frac{n_2}{n_1}$$

When the angle of incidence is i_c, the refracted ray travels along the surface of separation of the two media and there is said to be **grazing emergence** (Fig. 562 (i)). This phenomenon can be observed experimentally, and so also can the reverse effect of a ray travelling in medium (2) along the surface of refraction (**grazing incidence**) and entering medium (1) at an angle of i_c to the normal.

The question naturally arises as to what happens when a ray is incident

from medium (1) at an angle greater than i_c. The value of i_2 is always given by

$$\sin i_2 = \frac{n_1}{n_2} \cdot \sin i_1$$

But if $i_1 > i_c$, then $\sin i_1 > \sin i_c$, *i.e.*

$$\sin i_1 > \frac{n_2}{n_1} \quad \text{since } \sin i_c = \frac{n_2}{n_1}$$

$$\therefore \ \sin i_2 > \frac{n_1}{n_2} \cdot \frac{n_2}{n_1}$$

or

$$\sin i_2 > 1$$

There is no real angle whose sine exceeds unity, so that the information given by the law of refraction in a case in which the angle of

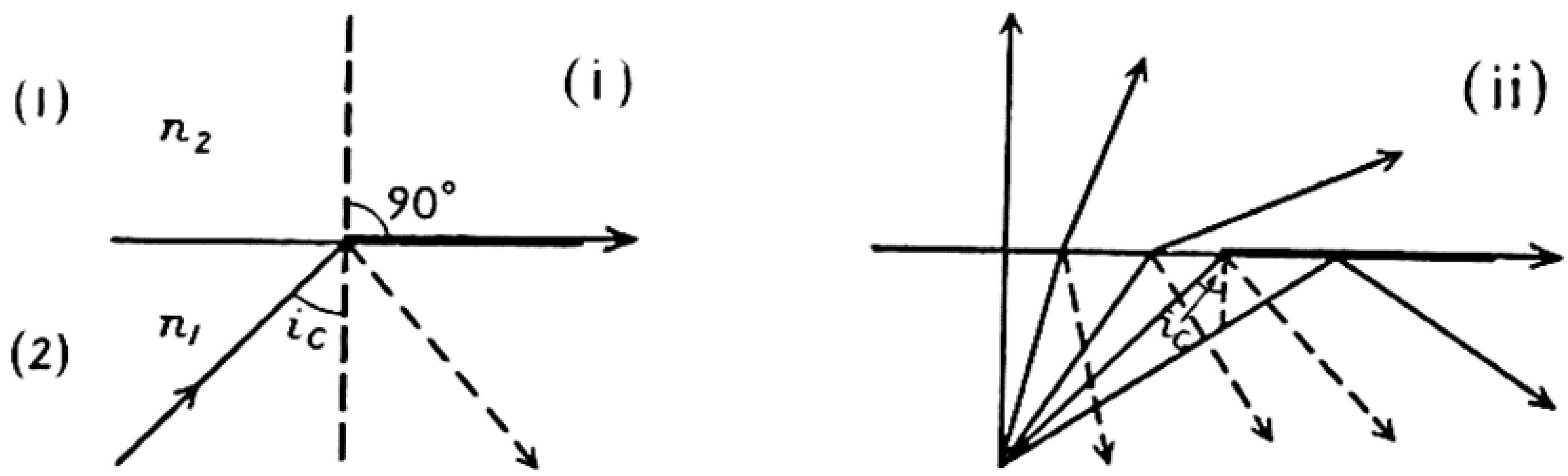


FIG. 562

incidence is greater than i_c is simply that refraction cannot take place. We remember now, however, that refraction is always accompanied by some reflection, and that the energy of the incident light is shared between the reflected and refracted rays. Therefore, when the angle of incidence exceeds i_c so that there is no refraction, all the light energy must be reflected and we have what is called **total reflection.** A boundary between two media therefore acts as a perfect reflector when light is incident on it from the medium of lower refractive index at an angle exceeding i_c, which is known as the **critical angle.** The sine of the critical angle is equal to the ratio of the smaller to the larger refractive index.

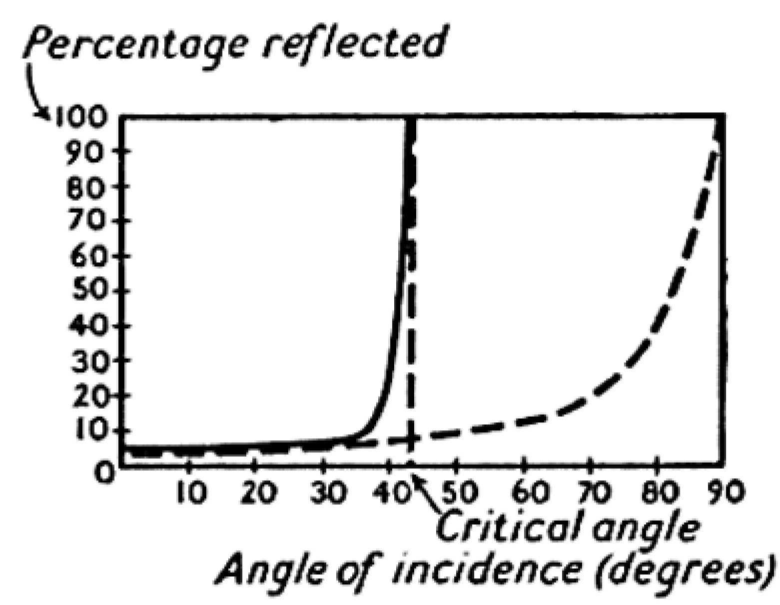


FIG. 563

Fig. 562 (ii) shows the transition from refraction and partial reflection to total reflection without refraction as the angle of incidence is increased.

In Fig. 563 the full line graph shows how the proportion of the incident light which is reflected varies with the angle of incidence for light passing

from glass to air. The rapid increase to perfect reflection at the critical angle should be noticed. The dotted graph shows the same thing for light passing from air to glass. A more gradual rise in the intensity of the reflected beam occurs as grazing incidence is approached.

With light passing from glass of refractive index 1·5 into air we have $n_1 = 1{\cdot}5$, $n_2 = 1$.

$$\therefore \sin i_c = \frac{n_2}{n_1} = \frac{1}{1{\cdot}5} = 0{\cdot}67 \text{ approximately}$$

Whence

$$i_c = 42° \text{ approximately}$$

For light travelling from glass to water we have $n_1 = 1{\cdot}5$, $n_2 = 1{\cdot}33$.

$$\therefore \sin i_c = \frac{n_2}{n_1} = \frac{1{\cdot}33}{1{\cdot}5} = 0{\cdot}89 \text{ approximately}$$

Whence

$$i_c = 63° \text{ approximately}$$

It is as well to remember that if the two faces of a parallel-sided slab of material are in contact with the same medium, *e.g.* a slab of glass in air or water, it is impossible for a ray which enters the slab through one face to be totally reflected at the opposite face. This is made clear in Fig. 564.

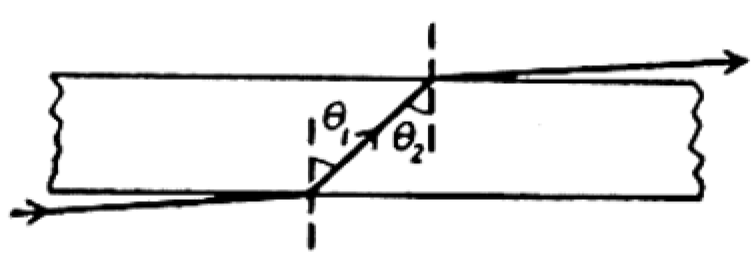


FIG. 564

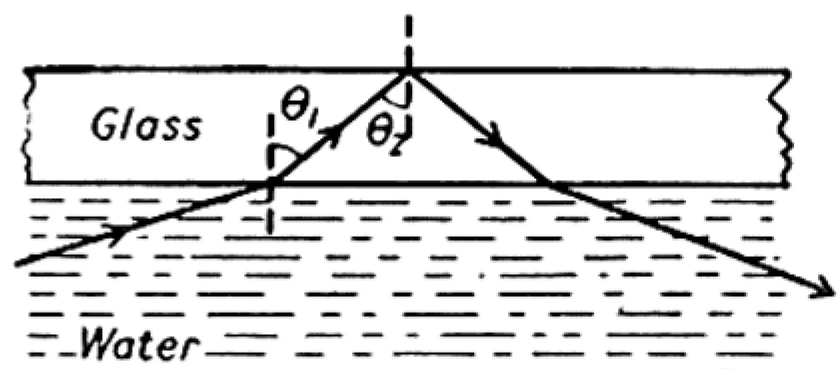


FIG. 565

The angle θ_1 has its maximum value (the critical angle) when there is grazing incidence, and θ_2 is always equal to θ_1. If however there is, say, water on one side and air on the other (Fig. 565), and light enters from the water, it is possible for θ_1 to have a value up to the critical angle for glass to water (63°). Thus θ_2 can have any value up to this, and when it exceeds 42° (the critical angle from glass to air) total reflection occurs at the second face. If an empty test-tube is held in a beaker of water its walls are seen to reflect light very strongly when viewed through the water at a suitable angle. This is due to total reflection at the *inner* wall of the test-tube, as illustrated in Fig. 565. It does not occur when water is placed in the test-tube because the conditions are then the same as in Fig. 564. Drops of water clinging to the inner walls of the test-tube appear as dark spots surrounded by bright areas where the reflection is occurring.

Applications of Total Reflection.—In many precision optical instruments in which reflection plays a part, total reflection is used in preference to reflection by an ordinary back-silvered mirror. The disadvantage of reflection by a sheet of glass which is silvered on the back is illustrated in Fig. 566. Some of the incident light is reflected from the outer unsilvered face of the glass (ray (1)), while the greater part of it is refracted into the glass, strongly reflected at the back-silvered face and then refracted again as it emerges (ray (2)). There is also some reflection of this ray back into the glass, and further reflection and refraction occurs as shown in the diagram, giving rise to the emergent rays (3), (4), etc.

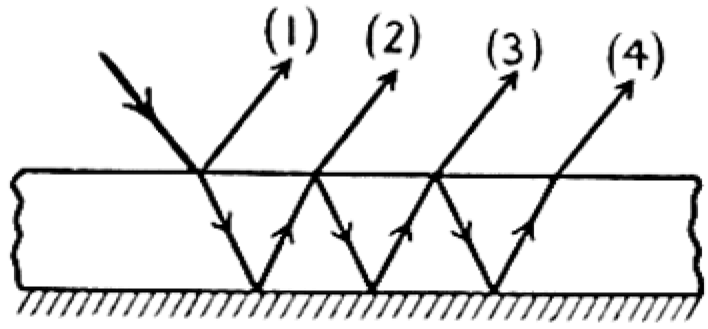


FIG. 566

Thus although ray (2) is usually by far the strongest ray on account of the strong reflection at the silvered surface, ray (1) is always present. Rays like (3) and (4), etc. are usually very weak, on account of absorption by the glass and the weak internal reflection at the front face.

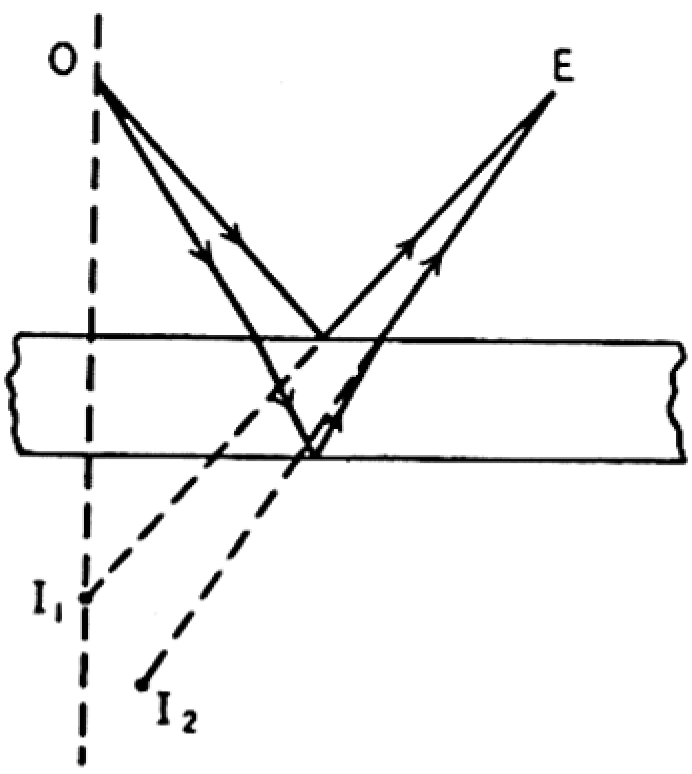


FIG. 567

An eye placed anywhere—say at E in Fig. 567—will receive rays of *all* the types (1), (2), (3), etc., although each one is due to a different incident ray coming from the object O. Those originating from one incident ray are parallel to each other and cannot all reach an eye placed at a given point. In order not to complicate the figure only types (1) and (2) are shown. The images due to these emergent rays occupy different positions because the rays entering the eye have different directions. The first image, I_1, lies on the normal through O, but I_2 and the other images lie on the side of the normal on which the eye is situated unless the angles of incidence are small, in which case all images are on the normal. The proportion of the incident light reflected by the front unsilvered surface becomes quite large when the eye is in such a position as to make the angle of incidence large (nearly grazing incidence), so that when this occurs I_1 may be the brightest image.

It is evident that in optical instruments like the periscope, range finder or binoculars, in which images are formed by lenses and reflectors, multiple images of the kind described above are a serious disadvantage. It might be thought that a remedy would be found by silvering the *front* of the mirrors instead of the back and thus avoiding the penetration of the light into the glass, but an exposed metal film of this kind is liable to oxydize and to become damaged when cleaned. The use of total reflection solves the problem admirably. Fig. 568 shows how a right-

angled glass prism can produce changes of 90° and 180° in the direction of light rays. The rays entering the prism are undeviated by refraction if they strike the first face normally, and simple geometry shows that the angle of incidence on the reflecting face is 45°. This ensures total reflection, because the critical angle from the prism glass to air is about 42°.

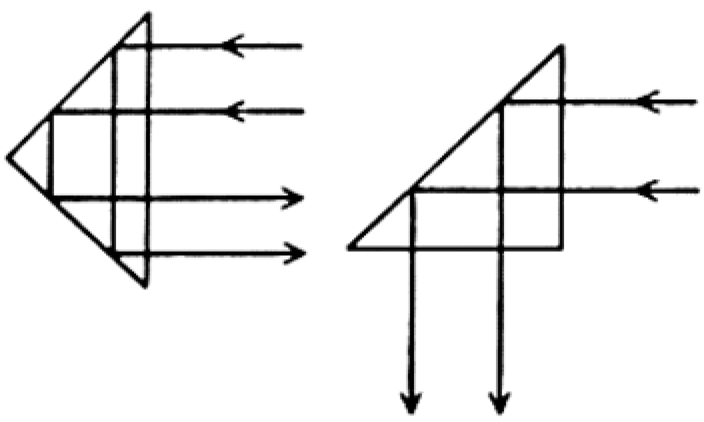


FIG. 568

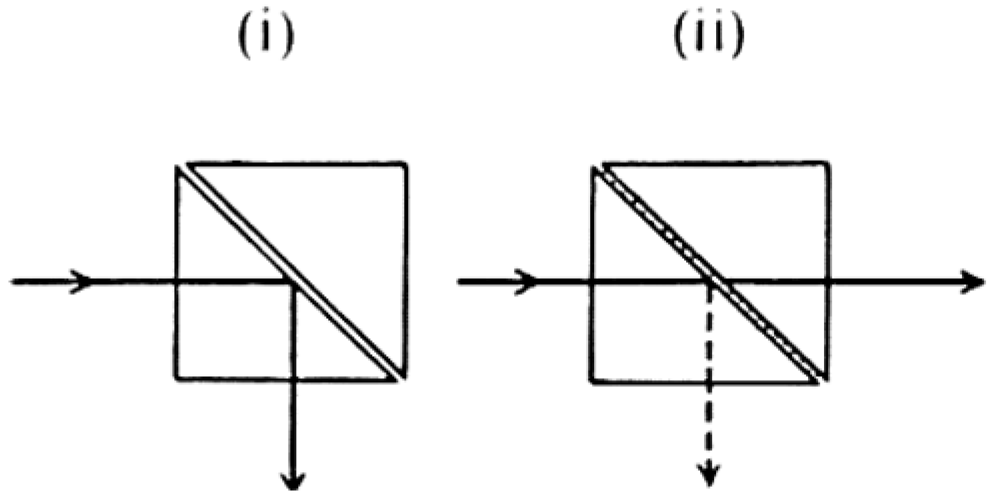


FIG. 569

A demonstration of total reflection involving the use of right-angled prisms can be made as follows. A beam of light (*e.g.* from a projection lantern) is sent normally into one of the small faces of one of a pair of 90° isosceles prisms arranged as shown (Fig. 569). There is a film of air between the two prisms in (i), and total reflection occurs causing a strong reflected beam. When a small quantity of water is introduced so as to form a thin film between the prisms (Fig. 569 (ii)), the critical angle is changed from 42° to 63° (page 750), and as the angle of incidence is 45°, the light is able to pass through and the reflected beam is correspondingly weakened.

There are several methods of determining refractive indices which are based on the measurement of the critical angle, and these are described later in this chapter.

4. REFRACTION BY A PRISM

We now consider the twofold refraction of a light ray which occurs when it passes through a piece of material via two plane faces which are inclined to each other (Fig. 570). We shall confine ourselves to cases in which the medium in contact with the plane surfaces is the same on both sides.

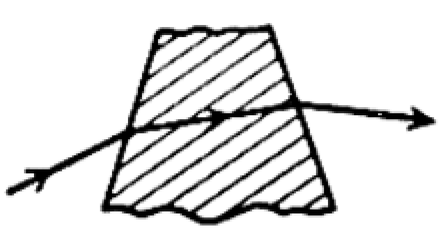


FIG. 570

The refracting material is usually in the form of a triangular prism whose triangular faces are perpendicular to the three rectangular faces (Fig. 571). The angle between the two faces at which refraction occurs is denoted by α and is called the **refracting angle,** while the line of intersection of the two faces is called the **refracting edge.** The path of the rays in a principal section of the prism is shown in Fig. 572, the plane containing the rays being perpendicular to both

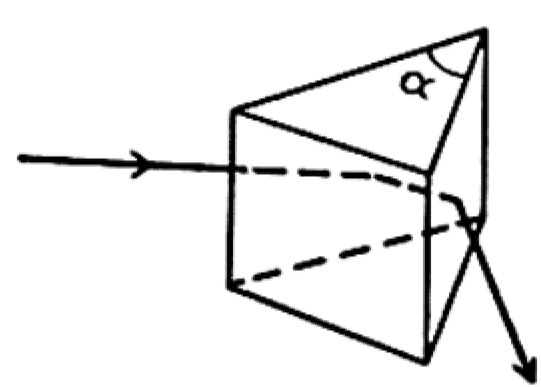


FIG. 571

the refracting faces. In this drawing the third side of the figure representing the section of the prism need not be put in, because the position and shape of this face does not affect the refraction which is under discussion.

Suppose that a ray from medium (1) is incident on the first face of the prism at B, the angle of incidence (i_1) being on the side of the normal further from the refracting edge A. The ray is refracted at an angle i_2 into the material of the prism (medium (2)), and is then incident on the second face of the prism at C at an angle i_3. Here the refraction causes the ray to emerge into medium (1) at an angle i_4 to the normal. It will be noticed that at each refraction the ray is bent away from the refracting edge.

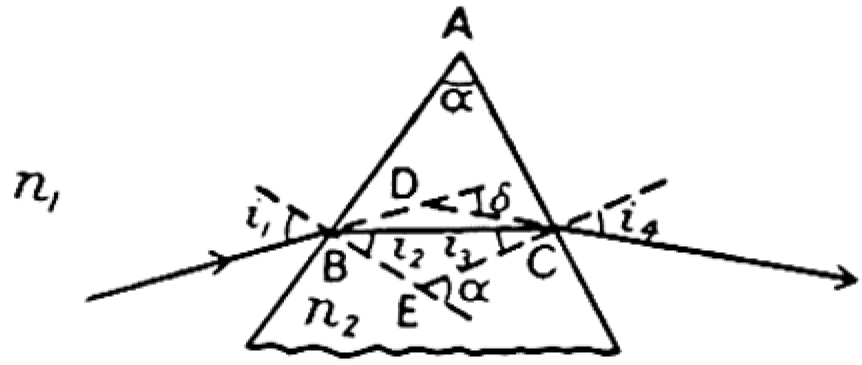


FIG. 572

If the incident and emergent rays are produced to meet at D, the angle between them, marked δ in the figure, is the deviation produced by the prism. The normals at B and C meet at E, and the angle between them is equal to the angle of the prism a. We now have

$$\widehat{DBE} = i_1 \quad \text{(vertically opposite angles)}$$

$$\therefore \; \widehat{DBC} = \widehat{DBE} - i_2$$

$$= i_1 - i_2$$

and similarly

$$\widehat{DCB} = i_4 - i_3$$

Also in triangle DBC

$$\delta = \widehat{DBC} + \widehat{DCB}$$

$$= i_1 - i_2 + i_4 - i_3$$

and in triangle EBC

$$a = i_2 + i_3$$

$$\therefore \; \delta = i_1 + i_4 - a$$

Minimum Deviation.—At this stage of the discussion it is important to remember that the path of the light through the prism is reversible. This means that a given deviation δ occurs for either of the two angles of incidence i_1 and i_4, which are related by the last equation. It is immaterial which face of the prism the light strikes first. If the light is sent into the left-hand face at an angle i_4 it will emerge from the opposite face at an angle i_1. Therefore, if a graph were plotted connecting the deviation δ with the angle of incidence, it would have such a shape as to show, in general, two values of the angle of incidence for each value of δ. It will be readily realized that a graph which has this property must exhibit

either a maximum or a minimum value of δ. Experiment easily shows that there is a minimum, and the graph has the shape shown in Fig. 573. It will be seen from the graph that as δ approaches its minimum value (δ_{min}) the two angles of incidence, i_1 and i_4, become more nearly equal and

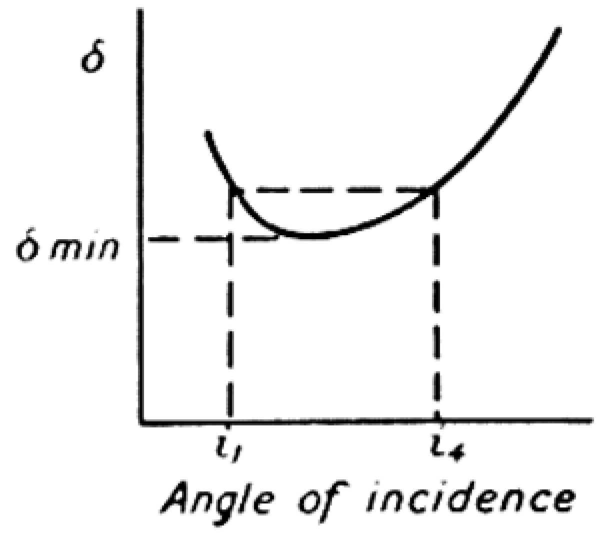


FIG. 573

are actually equal at the minimum. This means that for minimum deviation the ray must pass symmetrically through the prism, making the angles of incidence and emergence equal.

Thus when

$$\delta = \delta_{min}$$

$$i_1 = i_4$$

and since

$$n_1 \sin i_1 = n_2 \sin i_2$$

and

$$n_1 \sin i_4 = n_2 \sin i_3$$

therefore

$$i_2 = i_3$$

so that

$$\alpha = i_2 + i_3$$

$$= 2i_2$$

or

$$i_2 = \frac{\alpha}{2}$$

also

$$\delta_{min} = i_1 + i_4 - \alpha$$

$$= 2i_1 - \alpha$$

$$\therefore \ i_1 = \frac{\delta_{min} + \alpha}{2}$$

We have, therefore,

$$n_1 \sin\left(\frac{\delta_{min} + \alpha}{2}\right) = n_2 \sin\left(\frac{\alpha}{2}\right)$$

If, as is very frequently the case, medium (1) is air so that we can assume that $n_1 = 1$, and if n denotes the refractive index of the medium of the prism, then

$$n = n_2 = \frac{\sin\left(\dfrac{\delta_{min} + \alpha}{2}\right)}{\sin\left(\dfrac{\alpha}{2}\right)} \quad . \quad . \quad . \quad . \quad (3)$$

This is a very important equation, which enables us to determine the refractive index of the material of a prism with considerable accuracy, especially when the angles δ_{min} and α are measured by means of a spectrometer.

It is possible to plot the path of a ray through, say, a glass prism by using two pins to represent the direction of an incident ray and inserting two other pins on the other side of the prism to be in line with the first pair seen through the prism. The second pair marks the direction of the emergent ray and so the deviation can be measured. The graph of δ against angle of incidence can be constructed by repeating the above procedure for various suitably chosen angles of incidence. Each ray-trace provides two points on the graph, each having the same deviation but different angles of incidence (i_1 and i_4). It requires very careful insertion of the pins and measurement of the angles to obtain a good graph, because δ changes very slowly with the angle of incidence in the neighbourhood of the minimum. Nevertheless, for this very reason, the experiment can be made to yield a value for δ_{min} which is correct to one degree.

The refracting angle of the prism also can be measured optically by means of pins, one method being as follows. Two parallel lines are drawn on paper and marked with pins P_1, P_2 and P_4, P_5 (Fig. 574). The prism is then placed on the paper so that the incident rays represented by these lines fall on the two refracting faces. The images of P_1 and P_2 formed by *reflection* in the left-hand face are viewed, and the eye is moved until they are in line with each other. The pin P_3 is inserted to be in line with the two images. Similarly, P_6 is placed in line with the images of P_4 and P_5 seen in the other face. Lines representing the two faces are drawn, the prism is removed, and the holes made by P_3 and P_6 are joined to the points of incidence and the lines produced to meet at Q. The angle $P_3\widehat{Q}P_6$ is measured (let it be θ), and can be proved to be equal to 2α as follows. The larger angle (marked $360° - \theta$) between the two reflected rays can be regarded as the change of

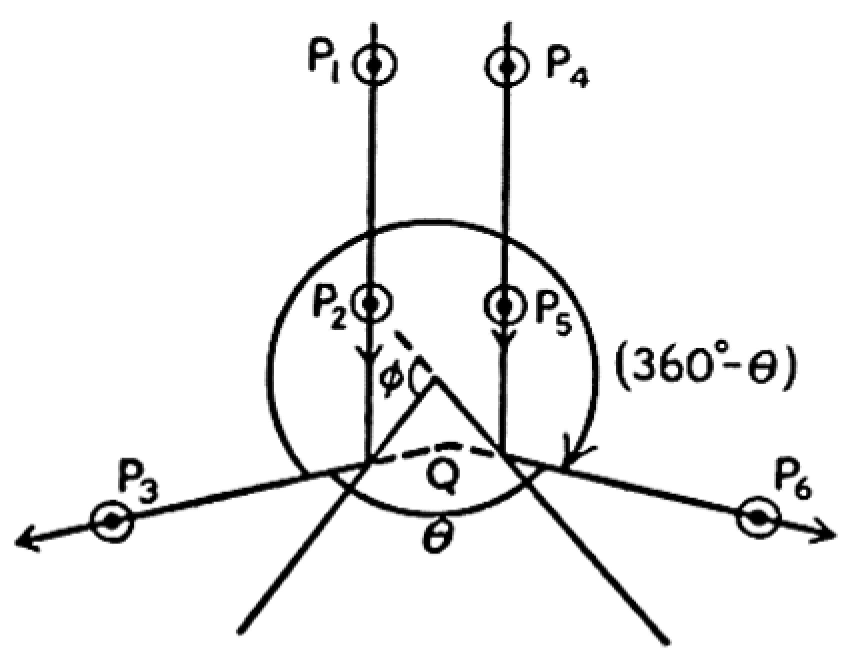


FIG. 574

direction of the reflected ray when the reflecting surface rotates through the angle marked ϕ.

$$\therefore\ 360° - \theta = 2\phi$$

But

$$\phi = 180° - \alpha$$

$$\therefore\ 360° - \theta = 360° - 2\alpha$$

$$\therefore\qquad \alpha = \frac{\theta}{2}$$

It will be noticed that the prism need not necessarily be placed symmetrically with respect to the two lines P_1P_2 and P_4P_5.

Refraction by a Thin Prism.—Suppose that the angle of incidence i_1 (Fig. 572) is small. This means that i_2 is also a small angle unless n_1 is very much greater than n_2, which is unlikely to be the case because the values of refractive indices of most substances lie between the limits 1 and 2.

If the refracting angle α is small—say 5° or less—the angle i_3 is also small, because α is the sum of i_2 and i_3. Finally, therefore, i_4 is also small. Consequently, we can replace the sines of the angles by their circular measure in the equations expressing the second law of refraction. Thus

$$n_1 i_1 = n_2 i_2$$

and

$$n_1 i_4 = n_2 i_3$$

We have therefore

$$\begin{aligned}\delta &= i_1 + i_4 - \alpha \\ &= \frac{n_2}{n_1}(i_2 + i_3) - \alpha \\ &= \frac{n_2}{n_1}\alpha - \alpha \quad \text{since } \alpha = i_2 + i_3 \\ &= \alpha\left(\frac{n_2}{n_1} - 1\right)\end{aligned}$$

If medium (1) is a vacuum (or air) and the material of the prism has a refractive index n, we have $n_1 = 1$, $n_2 = n$.

$$\therefore\ \delta = (n-1)\alpha$$

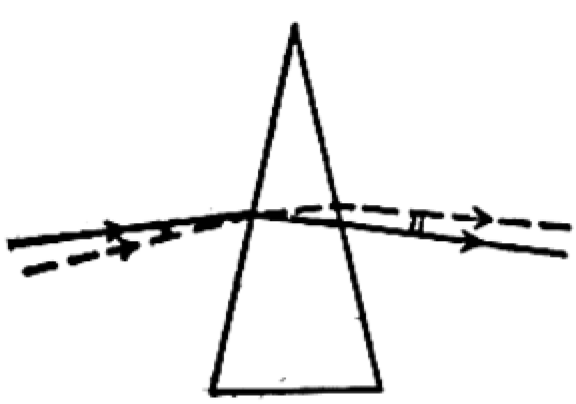


Fig. 575

This expression for the deviation produced by a thin prism or wedge of refracting material is valid only when the angle of incidence on the first face is small, and it will be noticed that under these conditions δ is independent of the angle of incidence. This means that if the direction of the incident ray is changed by a small amount, the emergent ray is rotated through the same angle (Fig. 575).

***Example.**—Find the largest refracting angle which a prism of material of refractive index* 1·662 *can have in order that light incident on one of its faces shall emerge from the opposite face.*

This question is answered by considering grazing incidence at the first face. The reason for this will be seen later. In Fig. 576 (i), the ray striking the left-hand face of the prism at an angle of 90° enters the prism at B. The angle of refraction will be the critical angle i_c. If the refracting angle of the prism is a, the angle of incidence on the second face is $(a - i_c)$. This angle must not exceed i_c if the ray is to emerge, the limiting case occurring when it is equal to i_c. Thus for emergence

$$(a - i_c) \text{ must not exceed } i_c$$

so that

$$a \text{ must not exceed } 2i_c$$

or, in the limiting case,

$$a = 2i_c$$

To show that this is the criterion required by the question, we must investigate the effect of allowing the original ray to strike the first face at B at an angle of incidence less than 90° (Fig. 576 (ii)). Let the first angle of refraction then be i, which will be less than i_c. The angle of incidence on the second face will then be $(a - i)$, which will be greater than the critical angle if a has its same value $(2i_c)$ as in the first case discussed. Hence the ray will not emerge from the second face of the prism unless a is reduced. Therefore the maximum possible value of a must be $2i_c$. But

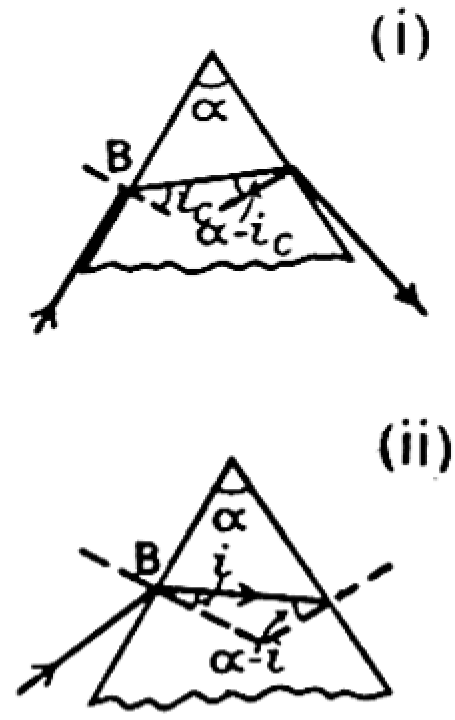


FIG. 576

$$\sin i_c = \frac{1}{1{\cdot}662}$$

$$= 0{\cdot}602$$

$$\therefore\ i_c = 37^\circ\ 0'$$

so that

$$a = 2(37^\circ\ 0')$$

$$= 74^\circ\ 0'$$

N.B.—For a refractive index of 1·5 the maximum possible refracting angle is about 83½°. Light cannot pass through a right-angle prism unless the refractive index is less than $\frac{1}{\sin 45^\circ}$, *i.e.* 1·414.

5. DETERMINATION OF REFRACTIVE INDEX OF SOLIDS

In this section we explain the principles of various methods of determining the refracting indices of solid media. In several cases the theory of the experiment has already been dealt with so that it is only necessary to mention the method for the sake of completeness.

Rectangular Block.—*Plotting Rays.* A rectangular block of, say, glass of refractive index n is laid flat on a horizontal sheet of paper and its outline drawn. Two pins, P_1 and P_2 (Fig. 577), are inserted, and another pair, P_3, P_4, fixed on the other side of the block so as to appear to be in line with the first pair when these are viewed through the block. When

the pins and block are removed, P_1P_2 and P_3P_4 are joined to represent the incident and emergent rays respectively. The points of incidence and emergence, A and B, are joined by a straight line which represents the path of the light through the block. Next the normals are drawn at A and B and the angles marked i_1, i_2, i_3 and i_4 measured. Then n can be found by evaluating $\frac{\sin i_1}{\sin i_2}$ and $\frac{\sin i_4}{\sin i_3}$. The two values should be equal within the limits of experimental error, and furthermore, i_1 should be equal to i_4, and i_2 to i_3. The determination can be repeated for various values of i_1.

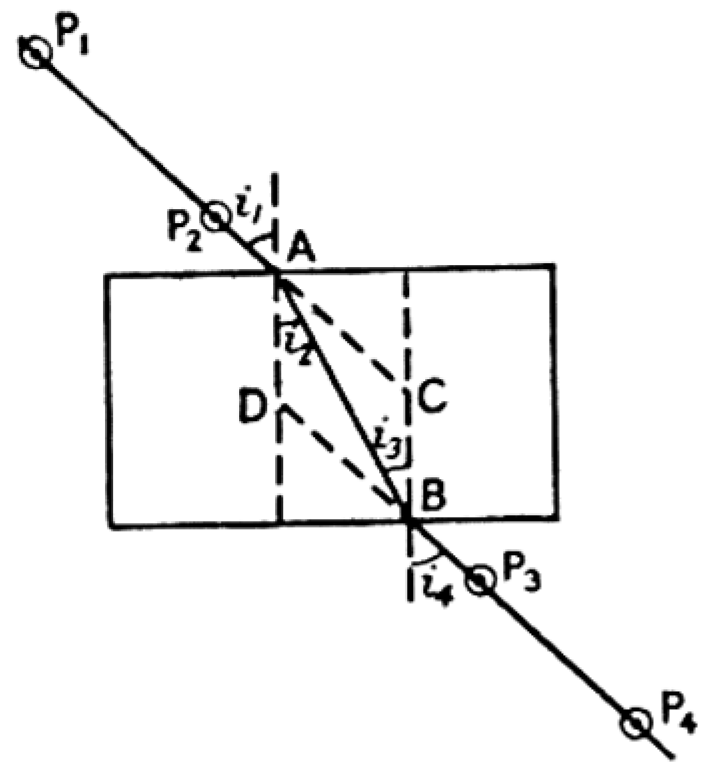


FIG. 577

A value of n can also be found by producing P_1P_2 forward to meet the normal at B in C. Then, as has already been shown (page 743),

$$n = \frac{AB}{AC}$$

Similarly, if P_3P_4 is produced backwards to D, we have

$$n = \frac{AB}{DB}$$

If a single pin is placed up against one face of the block and the direction of an emergent ray plotted from the other side, it is possible to obtain one pair of angles and one pair of lengths.

Real and Apparent Thickness. A single pin, P_0, is placed against one side of the block and viewed *normally* from the other side (Fig. 578). It appears to be nearer the eye than it actually is, and its apparent position P_0' can be located by moving a second pin P until this pin viewed directly appears to be in the same position as P_0 seen through the block. The adjustment is made by the elimination of parallax, the eye being moved slightly from side to side. Then the real thickness of the block is AP_0 and its apparent thickness is AP_0' (= AP), so that

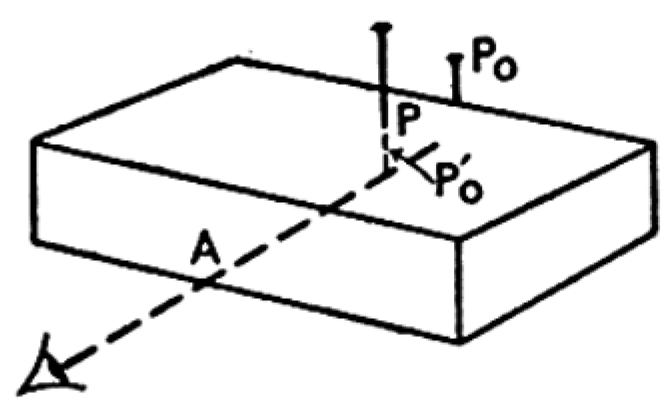


FIG. 578

$$n = \frac{AP_0}{AP}$$

This experiment is *not* capable of very great accuracy but is instructive nevertheless.

Triangular Prism.—We have already described the refraction of light through a triangular prism and the phenomenon of minimum deviation. One of the most accurate methods of determining refractive indices consists in the measurement of the minimum deviation and the angle of a prism of the material. These values are then substituted in equation (3). With a good spectrometer the results are accurate to the third place of decimals, *i.e.* to four significant figures. The variation of refractive index with the colour of the light can easily be detected and measured by means of a spectrometer. A fuller account of the construction and use of a spectrometer is given on page 962.

Critical-Angle Methods—Semicircular Slab.—This is the simplest experiment involving critical-angle determination. A semicircular slab (usually of glass) is laid on drawing-paper (Fig. 579) and its outline is drawn. A pin, P_0, is inserted at the centre of the flat rectangular face of the slab or else a permanent scratch is made on the face at this place. The experiment is greatly improved if vertical white screens are placed at SS as shown in the figure so as to reflect diffused light on to the flat surface of the block.

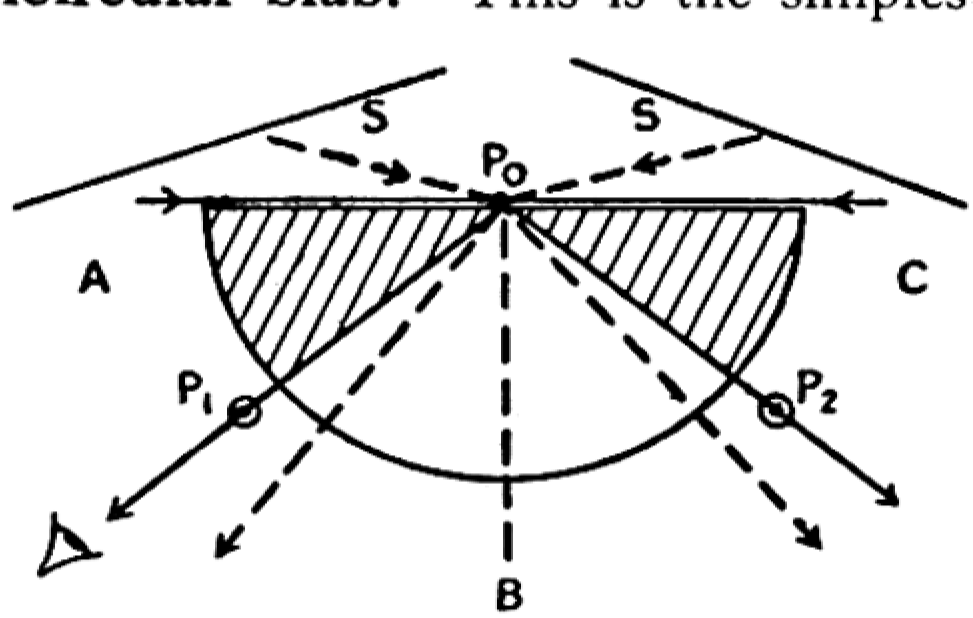


FIG. 579

On looking into the block through the curved surface and moving the eye round from A to B, the brightness of the flat face is seen to change. At first it is dark, and then at a certain position of the eye it becomes brighter near the end C. Further movement of the eye towards B causes the area of brighter illumination to spread towards A. The vertical line of division between the light and dark areas is not perfectly sharp if white light is used, because refractive index and therefore critical angle depends on the colour of the light. The eye is halted when the dividing line reaches the centre of the flat face at P_0, and the pin P_1 is inserted so as to be in line with P_0 and the eye when this stage is reached. If P_0 is a pin it will be noticed that it becomes invisible at this stage, but if it is a scratch it remains visible. A similar and symmetrical position is found between B and C and marked with the pin P_2, the brighter part of the surface being nearer to A in this case. The block is then removed and P_0 is joined to P_1 and to P_2.

Now, the brightness of the surface at the point P_0 suffers an abrupt change as an eye which is looking at P_0 moves across the line P_0P_1. This is due to the fact that on the side of P_0P_1 nearer B the eye receives light which enters the flat surface (from the screen) and is refracted as shown by the dotted rays. The line P_0P_1 is a refracted ray corresponding to a ray entering the material at P_0 at grazing incidence. If P_0B is the normal

at P_0, the angle $P_1\hat{P_0}B$ is the critical angle, and there can be no refracted rays coming from P_0 at a greater angle than this. This is the cause of the darkness at P_0 when the eye is moved off the line P_0P_1 towards A. Similarly, $P_2\hat{P_0}B$ is also the critical angle. It will be realized, of course, that for rays coming from P_0 there is no refraction at the curved surface. Thus the angle $P_1\hat{P_0}P_2$ is measured, and if n is the refractive index of the material of the block we have

$$n = \frac{1}{\sin\,(\textit{critical angle})} = \frac{1}{\sin\left(\dfrac{P_1\hat{P_0}P_2}{2}\right)}$$

6. DETERMINATION OF REFRACTIVE INDEX OF LIQUIDS

Hollow Prism.—Fig. 580 (i) is an illustration of a hollow prism. It is essential that the two pieces of glass forming the faces through which the light passes should each be of uniform thickness. Thus, if each is a

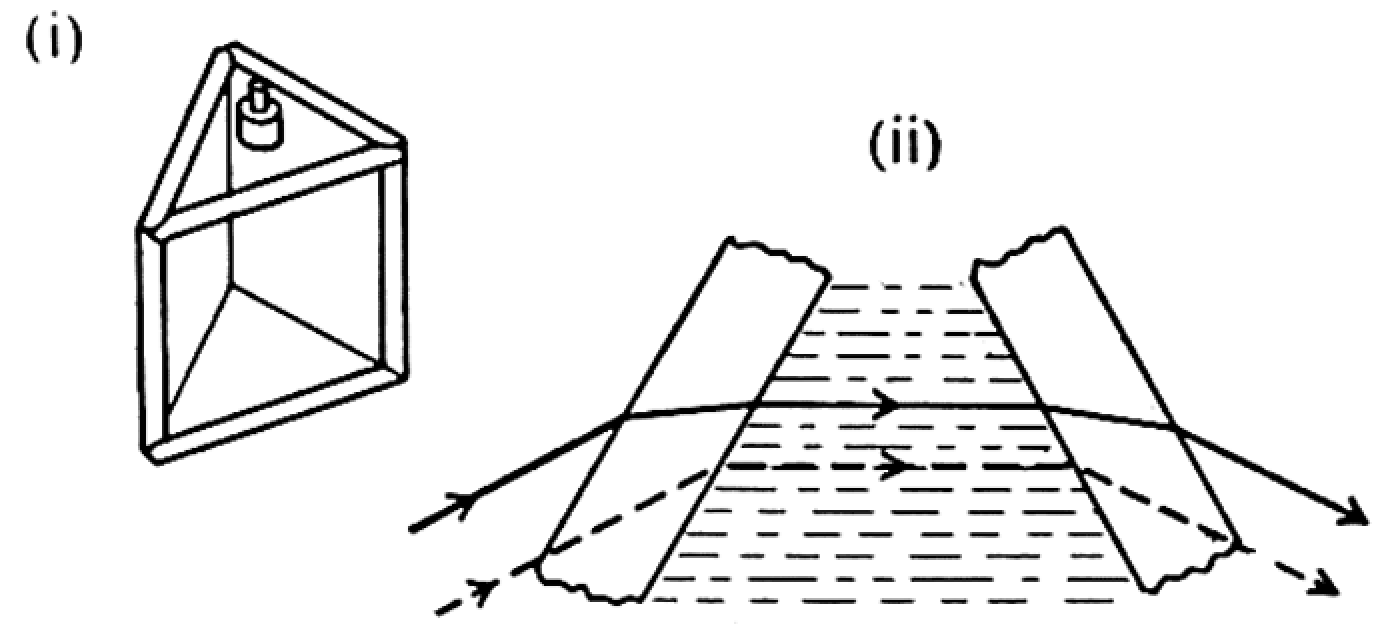


FIG. 580

parallel-sided sheet it does not influence the deviation of the light, which is, therefore, due entirely to the prism of liquid contained between the glass faces. The full lines representing the actual path of the light (Fig. 580 (ii)) are effectively the same as the dotted lines, corresponding rays in air and in the liquid being parallel to each other. Consequently the refractive index of the liquid can be found from measurements of the angle of the prism and minimum deviation exactly as if the prism were solid. For accurate work a spectrometer must be used.

Real and Apparent Depth.—The measurement of real and apparent depth and the use of the formula on page 744 constitutes a reasonably good method for the determination of the refractive index of a liquid. A common form of the experiment is shown in Fig. 581. The liquid is contained in a tall jar, at the bottom of which is a bright object O such as

a piece of metal. The actual depth (OA) of the liquid is easily determined. To find the apparent position (I) of the object when it is viewed by an eye situated vertically above it at E, a second bright object O′ is held above the liquid surface and it is moved up and down until its image formed by reflection in the surface of the liquid is at the same horizontal level as I. This condition can be judged fairly accurately by the parallax method. When the two images are at the same level they do not move relative to each other if the eye is moved horizontally. When the correct position of O′ has been found, the required apparent depth (IA) is equal to O′A,

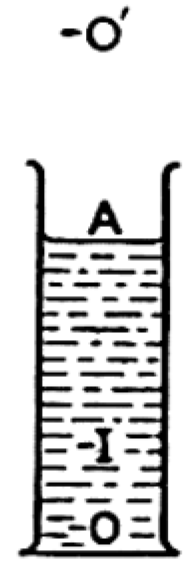


FIG. 581

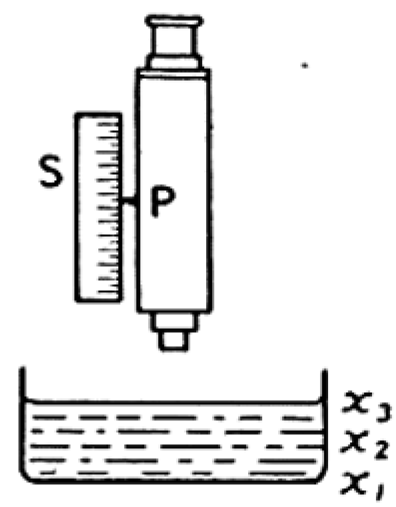


FIG. 582

which can be measured, because the image of O′ formed by reflection in the plane surface at A is as far below the surface as O′ is above it.

A travelling microscope is often used in a more precise form of the experiment requiring a smaller amount of liquid. This is shown in Fig. 582. The microscope must have a fairly small depth of focus, and it must be possible to move the microscope bodily up and down in the direction of its axis and to measure the amount of this movement on a suitable vertical linear scale S with a vernier or micrometer-screw refinement if possible. The microscope is mounted vertically above an empty dish and raised or lowered until a clear image of a scratch or other small mark on the inside of the bottom of the dish is seen. The final adjustment should be made by the elimination of parallax between the image and the crosswires as seen on looking through the microscope. Let the reading of the pointer P attached to the microscope be x_1 on the scale S. Some liquid is then put into the dish and the same mark on the bottom is sighted and parallax eliminated by raising the microscope (reading x_2). Finally, the microscope is raised until the top surface of the liquid is sighted (with the help of a little dust scattered on the surface) and the reading is x_3. The real depth is then given by $(x_1 - x_3)$ and the apparent depth by $(x_2 - x_3)$. This method can, of course, be used for a transparent solid slab as well as for a liquid.

Concave-Mirror Method.—This method is suitable for a small quantity of liquid. A concave mirror of radius of curvature about 20 cm. is laid face upwards on a horizontal surface (Fig. 583 (i)). A pin O held in a clamp above the mirror is moved up and down until it is on the same horizontal level as its own real image formed by the mirror. To make

this adjustment the eye must look down into the mirror from a position above O. The pin is then situated at the centre of curvature of the mirror (page 731). Next, liquid is poured into the mirror to a sufficient depth to ensure that the air/liquid surface is plane near its centre, and the height (OA) of O above the surface is measured. Then the pin is again adjusted until there is no parallax between itself and its image seen by looking vertically into the liquid-covered mirror. The new position of the pin (O′) will be below the original position (O), and the distance O′A is measured. When the pin is at O′ its image is formed (also at O′) by the rays shown in the diagram. Both the incident and the reflected rays are refracted on passing through the air/liquid surface. In the liquid, the rays must be normal to the mirror surface because they are reflected back along their own path. Therefore the rays in the liquid would, if produced upwards, meet at O. The angles i_1 and i_2 between the rays and the normal at B are marked in the enlarged drawing, and are clearly respectively equal to the angles $A\hat{O'}B$ and $A\hat{O}B$ in Fig. 583 (i). If the refractive index of the liquid is n, we have

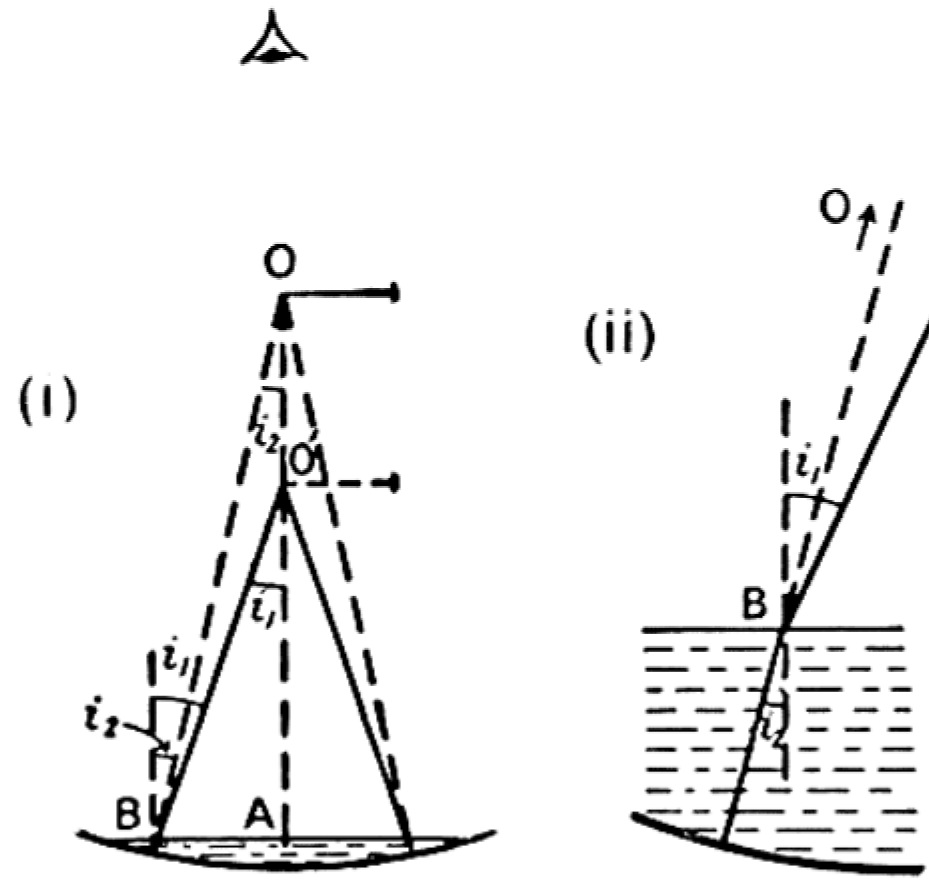


FIG. 583

$$n \sin i_2 = \sin i_1$$

or

$$n = \frac{\sin i_1}{\sin i_2}$$

$$= \frac{AB}{O'B} \div \frac{AB}{OB}$$

$$= \frac{OB}{O'B}$$

$$= \frac{OA}{O'A} \quad \text{approximately, if AB is small compared with OB and O}'\text{B}$$

Critical-Angle Methods.—*Wollaston's Method.* In this experiment a glass block G (Fig. 584), usually in the form of a cube, is placed on a few drops of liquid L resting on a dark surface such as a flat piece of ebonite. S is a diffuse source of light such as a ground-glass screen with a lamp behind it. The experiment is more satisfactory if the wave-length

range of the light is restricted by a coloured filter or by using a monochromatic source such as a sodium flame or lamp. There is a horizontal scratch on the glass block at A, and E is a slit in a screen which can be moved either up and down or towards and away from the block. The observer looks through the slit and receives light from a point on the

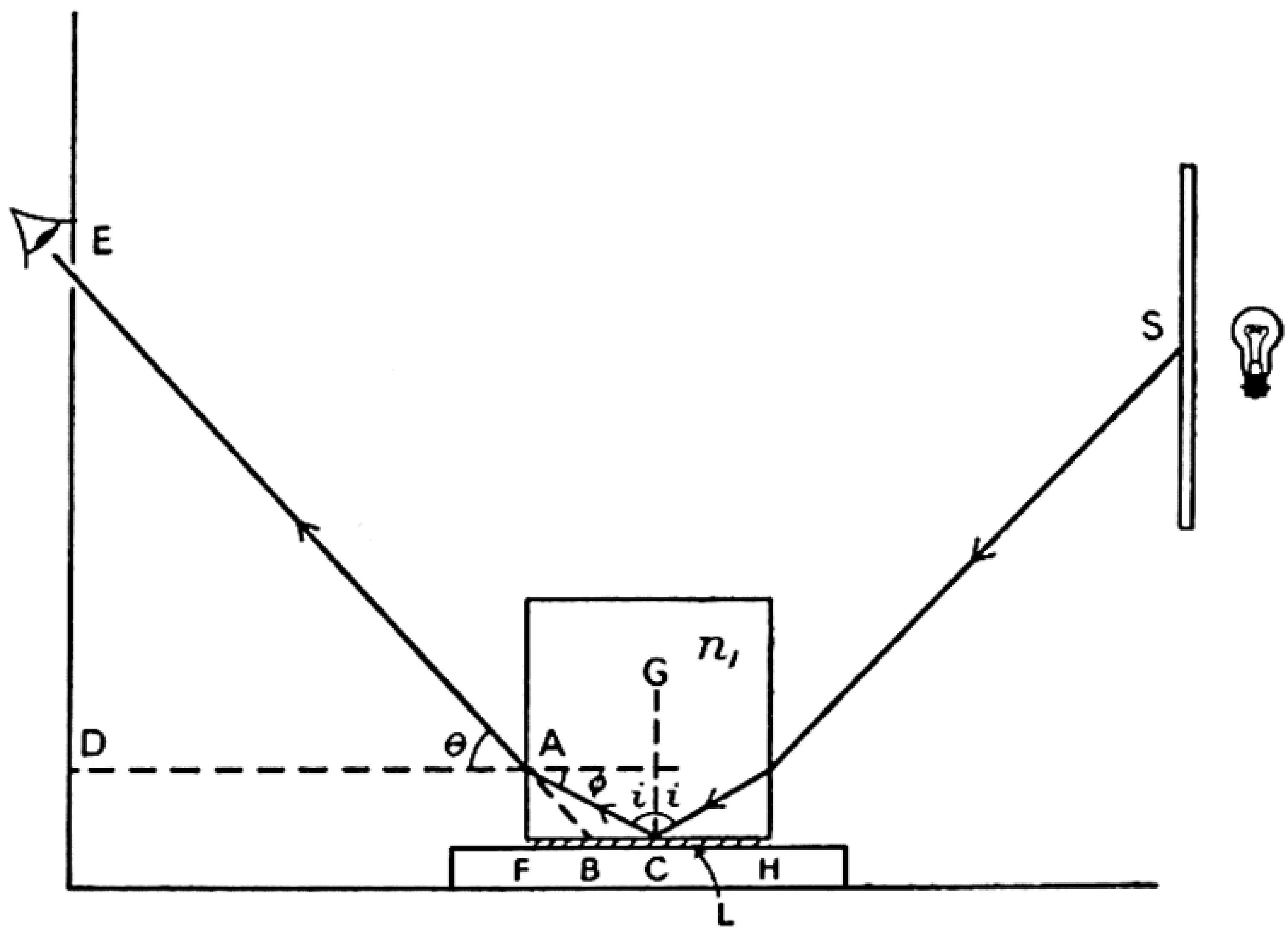


FIG. 584

illuminated screen by a path such as that indicated by the rays in the diagram. When E is high up or the screen is near the block (according to how its position can be altered) the angles marked i in the diagram are small, and if they are less than the critical angle for light going from glass to liquid the reflection at the glass/liquid surface is not total. As E is moved downwards the angles i approach the critical angle, and when this value is reached the reflection becomes total and the glass/liquid surface appears brighter. A line of division between the bright and less bright portions of the base of the block appears at this stage, and E is moved until this line appears to coincide with the scratch A. The light which enters the slit E then leaves the base of the cube at the critical angle (*i.e.* $i = i_c$) at C. If E is fixed in this position and the eye looks through it at the region FB of the base, the angle at which the light then entering the eye is reflected from the glass/liquid surface is less than the critical angle, so that the region FB appears less bright than the region BH in which the angle of reflection exceeds the critical angle.

When the setting has been made as described above, the angle θ can be determined, because its tangent is equal to $\dfrac{\text{ED}}{\text{DA}}$, where D is the point vertically below E on the same horizontal level as A. Then, by Snell's

law, the angle marked ϕ is given by

$$n_1 \sin \phi = \sin \theta \qquad (4)$$

where n_1 is the refractive index of the glass.

The angle i (which is the critical angle i_c) is equal to $(90° - \phi)$, so that

$$\cos i_c = \sin \phi$$

$$= \frac{\sin \theta}{n_1} \quad \text{by equation (4)}$$

and

$$\sin i_c = \sqrt{1 - \cos^2 i_c}$$

$$= \frac{\sqrt{n_1^2 - \sin^2 \theta}}{n_1} \qquad (5)$$

Furthermore, if n_2 is the refractive index of the liquid, we have, by Snell's law,

$$n_2 = n_1 \sin i_c$$

$$= \sqrt{n_1^2 - \sin^2 \theta} \quad \text{by equation (5)}$$

It is necessary to know the value of n_1, the refractive index of the material of the block. This may be obtained from the last equation by measuring θ, when a liquid of known refractive index (n_2) is present.

The **Pulfrich Refractometer** is a precision instrument in which the principle is similar to that of Wollaston's method and is shown in Fig. 585. The top of a block of glass is shaped as shown, its centre portion being highly polished. Round this portion a glass ring fits so as to allow liquid to be placed on the polished portion. Light from a monochromatic source S passes into the liquid and is refracted into the block, the greatest possible angle of refraction being i_c, the critical angle between the liquid and the glass. The light then emerges from the vertical face of the block, which is made accurately perpendicular to the top surface, and is received in a telescope which moves round a circular scale. There will, of course, be some light leaving the liquid/glass surface at an angle less than i_c and emerging from the block below the critical ray shown in the diagram, but there can be no light above the critical emergent ray. The critical ray

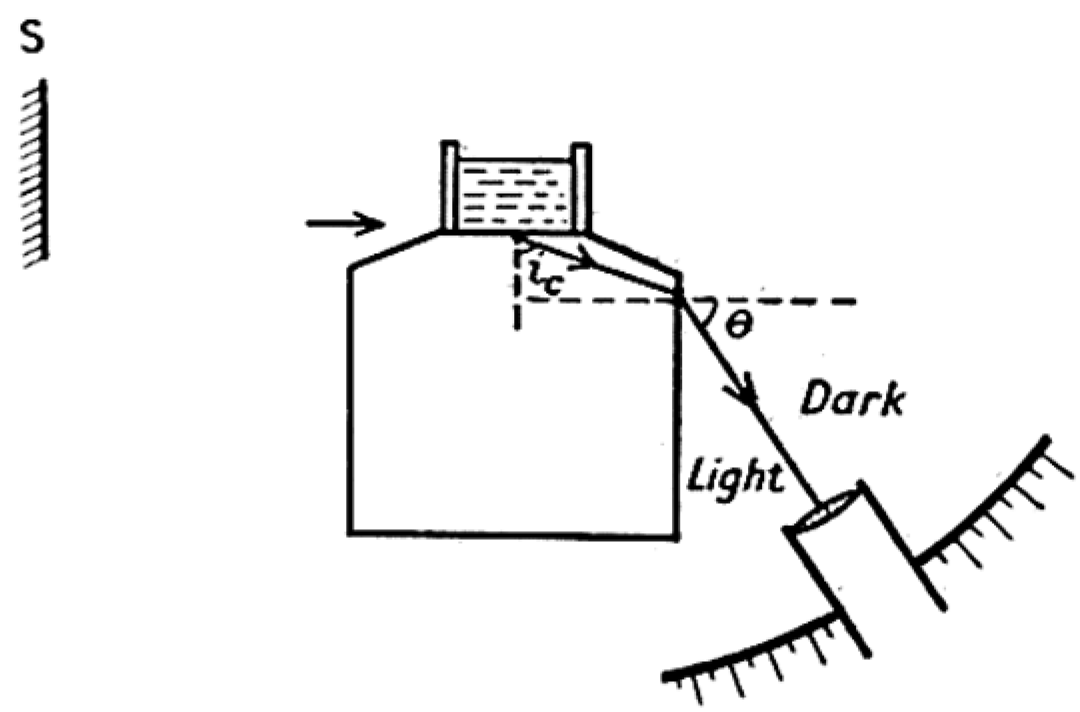


FIG. 585

can therefore be located by moving the telescope until the line of division between dark and light is on the crosswires. In this way θ can be found. The formula is the same as for Wollaston's method.

The Air-Cell Method.—This is quite a good teaching-laboratory method but it involves rather a large quantity of liquid. The air cell consists of a thin layer of air enclosed between two parallel glass plates separated by strips of tinfoil or other thin sheet round their edges. The edges of the plates are waxed to keep the plates together and to make the cell watertight. The cell is mounted on a vertical spindle and placed centrally in a vessel with plane glass sides as shown in Fig. 586 (i). In

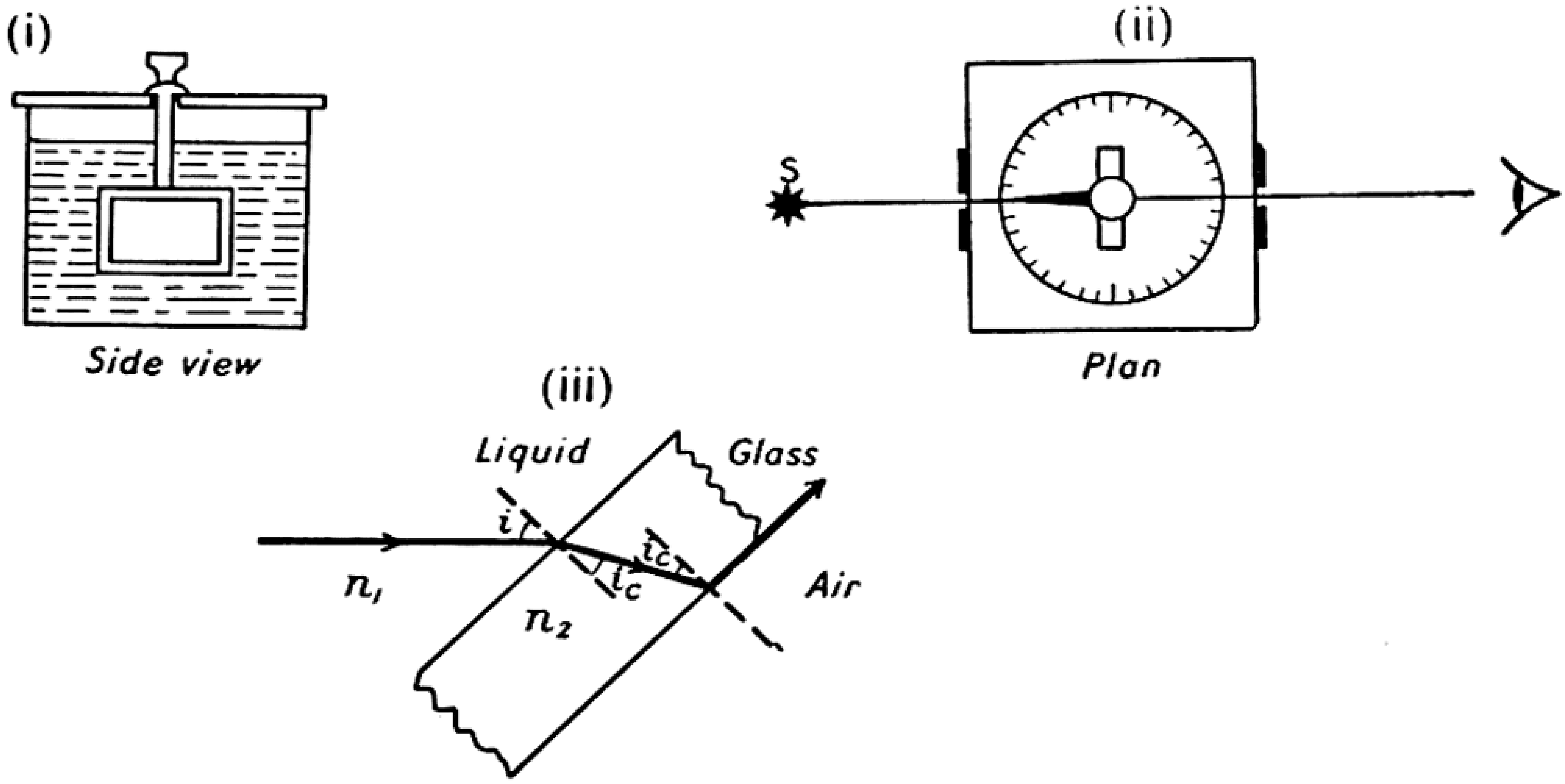


FIG. 586

the plan of the apparatus (Fig. 586 (ii)) two vertical slits are shown, one at each end of the tank, and a source of monochromatic light is placed as shown on the line joining the slits. A pointer is attached to the spindle and moves over a circular scale graduated in degrees. Liquid is placed in the tank, and the observer places his eye so as to receive the light from S through the two slits. The air cell is first turned until its plane is roughly perpendicular to the direction of the light. On rotating the cell in either direction from this position a setting will be reached at which the light is cut off, *i.e.* the far slit ceases to be visible to the observer. The pointer reading is taken and the cell is turned back through the perpendicular position until the slit again disappears. The pointer is read again, and the difference between the two readings is the angle through which the cell was turned between the two positions of extinction. The theory of the method is as follows.

The state of affairs in the critical condition is shown in Fig. 586 (iii). The critical reflection occurs at the glass/air surface, the critical angle being i_c. Since the glass plate has parallel sides, the angle of refraction

of the ray entering the glass is also i_C. If the angle of incidence from liquid to glass is i, and the refractive indices of the liquid and the glass are n_1 and n_2 respectively, we have, at the liquid/glass surface,

$$n_1 \sin i = n_2 \sin i_C$$

$$= 1 \quad \text{since } \sin i_C = \frac{1}{n_2}$$

Therefore

$$n_1 = \frac{1}{\sin i}$$

which means that in the critical case the angle of incidence at the liquid/glass surface is equal to the critical angle for light going from liquid to air. This angle is evidently half the angle through which the cell is turned between the two positions of extinction, so that n_1 can be calculated from the last equation.

The method is capable of improvement, *e.g.* by using parallel incident light and receiving it in a telescope.

7. DISPERSION

Dispersive Power.—It has already been mentioned that the refractive index of a given material varies with the wave-length of the radiation. This is known as **dispersion.** Normally, refractive index diminishes as wave-length increases, *i.e.* as the colour of the light changes through the visible region of the spectrum from violet to red. With ordinary glass of refractive index about 1·5, the values for violet and red differ by between 0·01 and 0·02. The principal facts concerning the formation of a spectrum by a prism of transparent material are mentioned on page 527 of Vol. 2, and an account of the study of spectra is included in Chapter LIII.

At this stage we shall merely discuss the effect of dispersion on the refraction of light by a thin prism. For the purpose of describing the performance of a prism and the optical properties of its material, three visible wave-lengths are chosen. These are usually a red line in the spectrum of hydrogen known as the C line and having a wave-length of approximately 6563 Å (Å = Ångström unit = 10^{-8} cm.), the yellow D line of the sodium spectrum of wave-length 5893 Å (actually there are two sodium lines differing in wave-length by about 6 Å, their mean being 5893 Å) and the blue F line of hydrogen (4861 Å).

Suppose that the values of the refractive index of a certain material are n_C, n_D and n_F respectively for the three specified wave-lengths. If light rays of these wave-lengths are simultaneously or separately allowed to fall nearly normally on one face of a thin prism of the material as shown in Fig. 587, then each colour will suffer a different deviation because the deviation δ is given by

$$\delta = (n-1)\alpha$$

where α is the refracting angle of the prism. Since $n_F > n_D > n_C$, therefore $\delta_F > \delta_D > \delta_C$, and if the incident rays all have the same direction, evidently the angle between the two extreme emergent rays is equal to

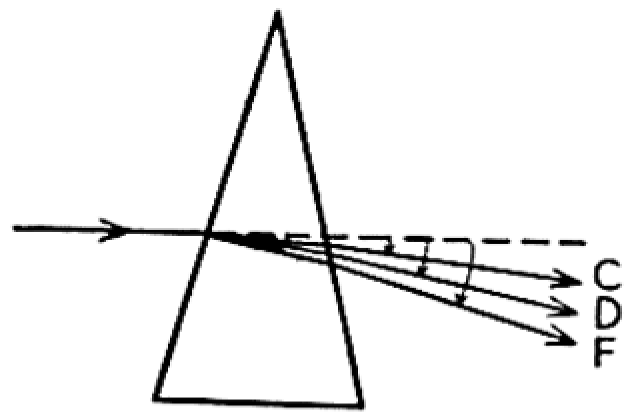


FIG. 587

$\delta_F - \delta_C$. This represents the dispersion due to the prism for the F and C wave-lengths, and is given by

$$\delta_F - \delta_C = \{(n_F - 1)\alpha\} - \{(n_C - 1)\alpha\}$$
$$= (n_F - n_C)\alpha$$

In order to specify the dispersive effect of the *material* of the prism we divide the dispersion due to the prism by the deviation of the intermediate wave-length—the yellow D light—thereby eliminating the refracting angle α and obtaining an expression which involves only the optical properties of the material. This expression is known as the **dispersive power** of the material for the specified wave-lengths, and we shall denote it by ω. Thus

$$\omega = \frac{\delta_F - \delta_C}{\delta_D} = \frac{(n_F - n_C)\alpha}{(n_D - 1)\alpha}$$
$$= \frac{n_F - n_C}{n_D - 1}.$$

The value of ω usually lies between 0·02 and 0·05.

Deviation without Dispersion.—Whenever white light is deviated by refraction, dispersion causes the separation of the constituent wave-lengths. This effect interferes very much with the formation of clear white-light images by prisms and lenses, the commonest fault from this cause being a coloration and consequent blurring of the edges of the images. This defect is called **chromatic aberration,** and is discussed later on (pages 825–833). The following account of how prisms may be used to cause deviation while the dispersion of any two wave-lengths is eliminated involves the basic principle by which chromatic aberration in lenses is minimized (page 827).

If a second prism identical with the first, but inverted, is placed so as to receive the emergent dispersed light (Fig. 588 (i)), it gives each ray an equal and opposite deviation to that which it received from the first prism, so that two rays such as C and F emerge from the combination

parallel to each other and to their initial direction. A converging lens placed so as to receive the emergent rays C and F would bring them to a common point-focus which would not be coloured. It would, in fact, be an image of the point A. The same recombination of the colours would occur if the two rays were received by an eye. In this sense, therefore, the effect of dispersion has been eliminated, but so also has the overall deviation—in fact the two prisms together are equivalent to a parallel-sided slab of material.

It would appear at first sight that a band of colour would be formed on a screen if it were placed so as to intercept the emergent beam in Fig. 588 (i)

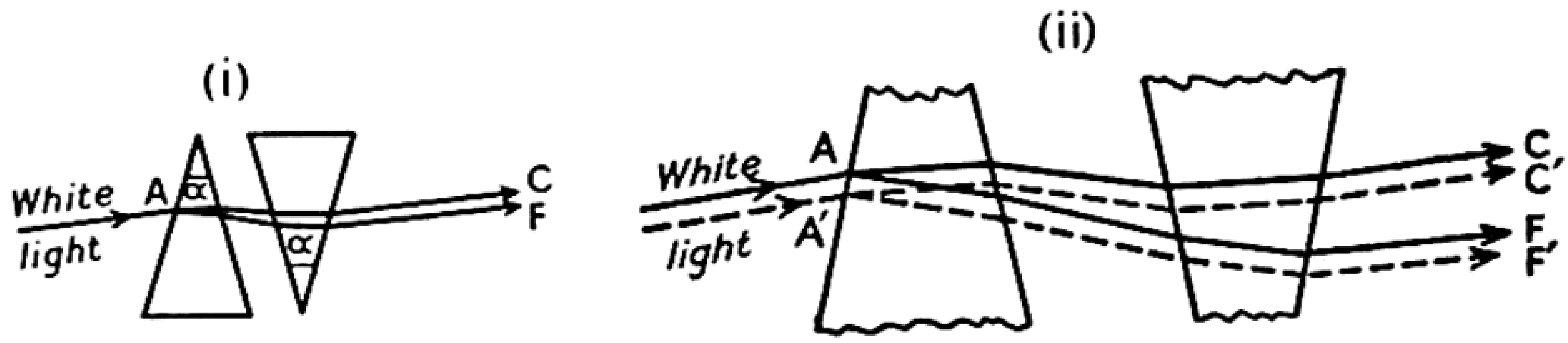


FIG. 588

without the interposition of a lens, because the parallel beam has different colours at different places across the width. It must be realized, however, that this apparently clear-cut sorting of the colours is due to our having supposed that a *single* white-light ray is incident on the first face of the first prism, that is to say, that all the colours present in the white light enter the prism at one and the same point. Since neither a point source nor a single ray can be obtained, the state of affairs is always more complicated in practice than is shown in Fig. 588 (i). In Fig. 588 (ii) a second white-light ray, parallel to the first, is shown (dotted) entering the first prism at A′ near A. This ray gives rise to an emergent parallel beam, bounded by the dotted rays C′ and F′ which, because of the finite widths of the beams, will overlap the first beam and there will be superposition of two spectral colours at any point in the region of overlap, *i.e.* between the rays C′ and F. An extension of this argument to the innumerable rays of which an incident beam of finite width can be supposed to be made up shows that the resultant emergent beam will consist of a mixture of spectral colours, *i.e.* white light, except at the edges. This state of affairs can be demonstrated by placing a screen in the path of the emergent beam. The two prisms can, of course, be replaced by a single parallel-sided slab, and the foregoing discussion applies equally well to this case. In the case shown in Fig. 588 (ii) the top of the emergent beam will be red and the bottom blue. The colours will be reversed if the white light enters the first prism on the other side of the normal.

If two prisms are made of different materials it is quite possible to give them such refracting angles that the dispersion due to one prism cancels that due to the other while the two deviations do not cancel. Thus, for

the first prism the angle between the emergent F and C rays is $(\delta_F - \delta_C)$ and is given by

$$\delta_F - \delta_C = (n_F - n_C)\alpha$$

while for the second prism

$$\delta_F' - \delta_C' = (n_F' - n_C')\alpha'$$

In order that the two dispersions shall cancel each other,

$$\delta_F - \delta_C = \delta_F' - \delta_C'$$

so that

$$(n_F - n_C)\alpha = (n_F' - n_C')\alpha' \quad . \quad . \quad . \quad . \quad (6)$$

Thus the value of α' which, for a given value of α, gives no dispersion, can be calculated from this equation, and when two prisms satisfying this equation are placed together in opposition they constitute an "achromatic pair" (Fig. 589), giving no dispersion of the C and F rays (although not necessarily eliminating it for any other pair of wavelengths). For the particular case of two prisms of the same material the equation gives $\alpha = \alpha'$, which, as mentioned above, eliminates deviation as well, but, in general, a certain amount of deviation remains when the materials are dissimilar.

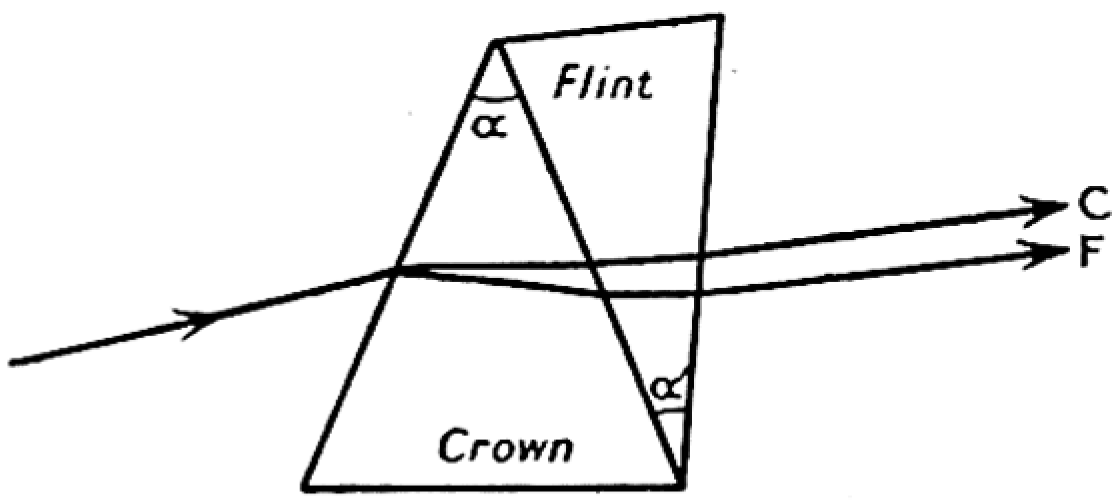


FIG. 589

Let us consider the deviation of the intermediate (D) ray for an achromatic pair of prisms. For the first prism

$$\delta_D = (n_D - 1)\alpha$$

so that

$$\alpha = \frac{\delta_D}{n_D - 1}$$

and for the second prism

$$\alpha' = \frac{\delta_D'}{n_D' - 1}$$

Since we are supposing that the pair of prisms is achromatic equation (6) is true, and if we substitute the above expressions for α and α' in equation (6) we obtain

$$\delta_D . \frac{n_F - n_C}{n_D - 1} = \delta_D' . \frac{n_F' - n_C'}{n_D' - 1}$$

or

$$\delta_D \omega = \delta_D' \omega' \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where ω and ω' are the dispersive powers of the two materials for the F and C lines.

Equation (7), therefore, shows that if the dispersive powers of the two materials were equal, δ_D and δ_D' would be equal, and the choice of a and a' which eliminated dispersion (*i.e.* which satisfied equation (6)) would also eliminate the deviation (which, for the D light, is equal to $(\delta_D - \delta_D')$). On the other hand, equation (7) shows that if the two materials have different dispersive powers, then $(\delta_D - \delta_D')$ will not be zero for an achromatic pair.

Furthermore, by using two different materials it is possible to have two prisms which produce no deviation. The relation between the angles for this condition is

$$(n_D - 1)a = (n_D' - 1)a' \qquad (8)$$

and, at the same time, dispersion is not eliminated because values of a and a' which satisfy equation (8) do not necessarily satisfy (7).

This principle has been used in an instrument called the **direct-vision spectroscope.** A form of this consists of two flint-glass prisms placed alternately between three crown-glass prisms, the two sets of prisms being inverted with respect to each other. Their angles are chosen so that yellow light is undeviated when it passes through the train of prisms. When parallel light from an illuminated slit enters the eye after passing through all the prisms, a spectrum is seen owing to the dispersion. The instrument is normally used only for rapid examination of spectra.

An achromatic pair of thin prisms is not a very important device in itself, but the principle of using two different materials to eliminate dispersion while leaving some deviation is very important in lens design. The two types of glass frequently used are crown glass and flint glass. The values of the relevant optical constants of two examples of these glasses are as follows:—

	n_F	n_D	n_C	$\omega\left(=\dfrac{n_F - n_C}{n_D - 1}\right)$
Crown glass . .	1·5230	1·5170	1·5145	0·0164
Flint glass . .	1·6637	1·6499	1·6434	0·0312

If, in equation (6), we take the plain letters to refer to crown glass and the dashed letters to flint glass, the condition for achromatism of a pair of prisms of these materials is

$$(1{\cdot}5230 - 1{\cdot}5145)a = (1{\cdot}6637 - 1{\cdot}6434)a'$$

or

$$\frac{a}{a'} = \frac{0{\cdot}0203}{0{\cdot}0085} = 2{\cdot}39$$

It should be noted that the achromatism produced by *two* prisms is only approximate inasmuch as only *two* colours are superimposed perfectly. For three-colour superposition—and therefore greater approximation to white light—three different prisms would be necessary.

EXAMPLES XLIV

1. A point source is viewed normally through a parallel-faced slab of glass thickness t and refractive index μ. Show that the displacement of the image is independent of the position of the source and find an expression for this displacement. Describe briefly the determination of the refractive index of glass by this method.

The lower half of a gas jar, 30 cm. tall, contains water and the upper half paraffin, of refractive indices 1·33 and 1·47 respectively. What is the apparent depth of the jar when viewed normally? (L.I.)

2. What is meant by *total reflection*? Two plane parallel pieces of glass are cemented together to enclose a layer of air between them. Describe the use of this apparatus to determine the refractive index of a transparent liquid and give the theory of the method.

Derive an expression for the deviation D, of a ray of light incident at an angle i on a transparent sphere and reflected once internally before emerging, in terms of i and the angle of refraction r. If the angle of incidence i_c corresponding to minimum deviation is given by $\cos i_c = \sqrt{\dfrac{\mu^2 - 1}{3}}$, determine the corresponding deviation for a water-drop of refractive index $\mu = 1{\cdot}331$. (L.A.)

3. Explain, with the help of a diagram, the term *critical angle*, and state the condition necessary for its existence.

Calculate the magnitude of the critical angle when the refractive indices of the two media concerned are 1·52 and 1·47 respectively.

Explain *two* practical applications of total reflection. (L.I.)

4. Find an expression connecting the refractive index of a glass prism with its refracting angle and the angle of minimum deviation.

ABC is a right-angled glass prism, the angles at A and C being 30° and 60° respectively. The face AB is silvered, and a ray of light incident on the face AC at an angle of 55° retraces its path after reflection from it. What is the refractive index of the glass? What would be the deviation of a ray incident at the same angle on the face BC and so directed that it emerges from AC without undergoing reflection at AB? (L.I.)

5. Explain the phenomenon of total internal reflection, and obtain the relation between refractive index and critical angle.

The face BC of a 60° glass prism, ABC, is coated with a liquid of refractive index 1·3. If the refractive index of the glass is 1·6, find the angle of incidence of a ray on AB in order that it shall strike BC at the critical angle.

What are the reasons for the use of right-angled prisms in binoculars? (L.I.)

6. What is meant by total internal reflection? Describe how this phenomenon may be demonstrated and state the conditions under which it occurs.

Derive a formula for the deviation of a ray of light by a wedge of transparent material of small refracting angle.

A ray of light, incident on one face of a wedge of refractive index 1·50 and angle 5°, is just totally reflected at the second face. Calculate the angle of incidence. (C.H.S.)

7. State the laws of refraction of light, and explain the meaning of *total reflection*.

ABC is a prism of refractive index 1·3, the angle at B being 90°. At what angle must a light ray strike the face AB in order that it may just emerge from the face BC? (L.I.)

8. A pencil of white light falls obliquely on a plane surface separating air and water. Draw and label simple ray diagrams of the optical phenomena observable and state the optical laws involved.

Calculate the deviation of a parallel pencil of monochromatic light, incident normally on one face of a 30° prism of refractive index 1·5, in passing through the prism. (L.I.)

9. What is total internal reflection? How may it be used to measure the refractive index of a liquid?

What must be the least angle of a prism of glass $n = 1{\cdot}50$ in order that no light can be refracted through it?

Trace the path of a ray entering one of the smaller faces of a 45°, 45°, 90° prism in a direction parallel to the base and perpendicular to the edges. (L.I.)

10. Describe a method, based on grazing incidence or total internal reflection, for finding the refractive index of water for the yellow light emitted by a sodium flame.

The refractive index of carbon bisulphide for red light is 1·634, and the difference between the critical angles for red and blue light at a carbon bisulphide-air interface is 0° 56′. What is the refractive index of carbon bisulphide for blue light? (J.M.B.H.S.)

11. Explain the term *critical angle*.

Describe and explain the air-cell method of finding the refractive index of a liquid. (L.I.)

12. Explain what happens in general when a ray of light strikes the surface separating transparent media such as water and glass.

Explain the circumstances in which total reflection occurs, and show how the critical angle is related to the refractive indices of the media.

Describe a method for determining the refractive index of a medium by means of critical reflection. (L.I.)

13. State the laws of refraction of light and describe a simple experiment to verify them.

A flat circular disc of glass is placed on a sheet of drawing-paper and the paths of rays incident on it are plotted in the usual way with pins. Part of the light in an incident ray emerges without internal reflection, being deviated through an angle D, while part emerges after one internal reflection. It is observed that when this latter part is parallel to the incident ray $D = 80°$. What is the refractive index of the glass? (L.I.)

14. How would you determine the refractive index of glass, using a glass disc of circular section about ½ in. high and 3 in. in diameter? A ray incident on the curved surface at an angle of 60° is observed to be deviated through 54° when it emerges from the disc. What is the refractive index of the glass? For what angle of incidence will the ray which emerges from this disc after one internal reflection be parallel to the incident ray though travelling in the opposite direction? (It is assumed throughout that the plane of incidence is at right angles to the axis of the disc.) (L.I.)

15. Define *refractive index*.

A ray of white light is incident normally on one face of a prism of refracting angle 30°. Calculate the angle between the violet and red rays which emerge from the opposite face if the refractive indices of the glass for violet and red light are 1·630 and 1·614 respectively.

Draw a diagram showing the paths of the rays through the prism. (L.I.)

16. Derive the formula for the deviation produced when a ray of light is incident nearly normally on one surface of a triangular prism of small refracting angle and passes unsymmetrically, but in a principal plane, through the prism.

A prism of glass of refractive index 1·62 has a refracting angle of 5°. What is the smallest angle of a second prism of glass of refractive index 1·52 which, when combined with the first, will produce a total deviation of (*a*) ½°, (*b*) 6°? (L.Med.)

17. State the laws of refraction of light, and explain what is meant by *dispersion*.

A ray of white light in air is incident normally on one face of a flint-glass prism. It traverses the glass and emerges into the air through another face which is inclined at 30° to the former. If the refractive indices of the glass for blue and red light are 1·664 and 1·644 respectively, calculate the angular separation of these coloured rays after emerging from the prism. (L.I.)

18. Find an expression for the deviation of a ray of light which passes nearly normally through a thin prism of refracting angle α, made of material of refractive index μ.

Find the angle of a thin prism made of flint glass which can be combined with a thin prism of angle 2° made of soda glass so that there is no resultant deviation of a yellow ray passing through the system, the refractive indices of the two glasses for light of this colour being $1{\cdot}6$ and $1{\cdot}5$ respectively.

Explain the action of the train of prisms in a direct vision spectroscope by means of a simple ray diagram. (L.Med.)

Chapter XLV

REFRACTION BY SPHERICAL SURFACES

1. REFRACTION AT A SINGLE SPHERICAL SURFACE

Suppose that the surface between two transparent media ((1) and (2)) of refractive indices n_1 and n_2 is part of a sphere, the centre of curvature of which is C (Fig. 590). P is the central point of the refracting surface and a point object O is situated on the axis CP. A ray from O is incident on the refracting surface at A, the angle between the ray and the prolongation of the normal CA being i_1. The angle of refraction is i_2, and since this is

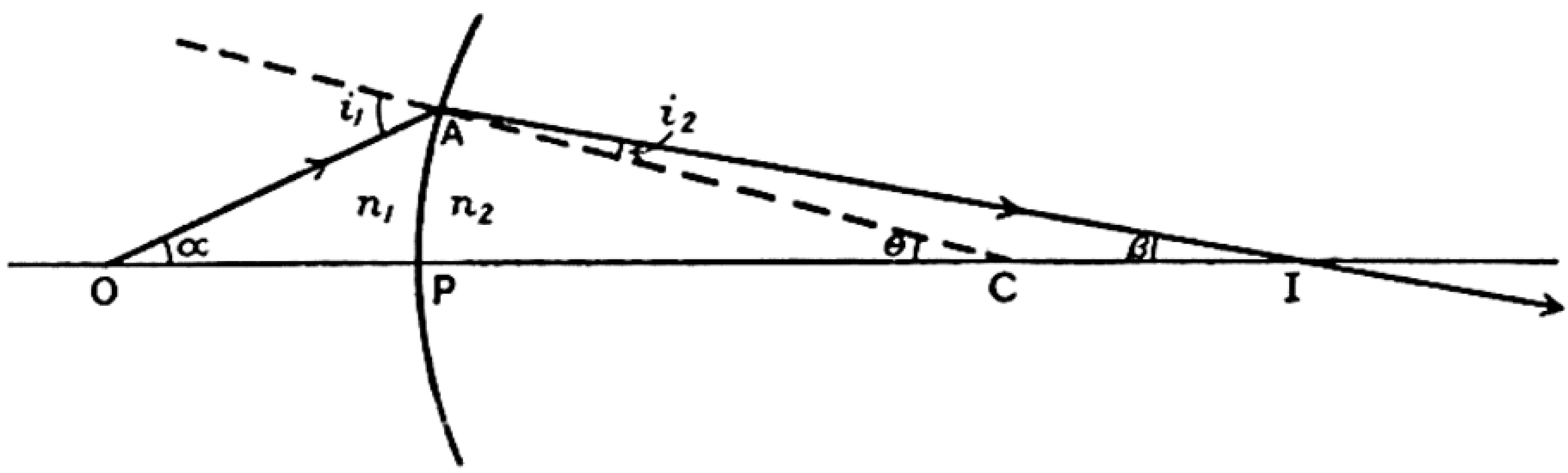


FIG. 590

assumed to be smaller than i_1, we are supposing that n_2 is greater than n_1. It is also assumed that the deviation $(i_1 - i_2)$ is sufficient for the refracted ray to converge towards the axis and to cross it at I.

In triangle AOC, of which i_1 is an exterior angle, we have

$$i_1 = \alpha + \theta$$

and for triangle ACI

$$\theta = i_2 + \beta$$

so that

$$i_2 = \theta - \beta$$

We now suppose that AP is a short distance compared with the distances of O, C and I from P. This makes i_1 and i_2 small angles whose sines can be put equal to their circular measure. Snell's law can therefore be expressed as

$$n_1 i_1 = n_2 i_2$$

or

$$n_1(\alpha + \theta) = n_2(\theta - \beta)$$

i.e.

$$n_2\beta + n_1\alpha = (n_2 - n_1)\theta$$

Next, as in the case of the spherical mirror, we can, since the angles are small, put

$$\alpha = \frac{AP}{OP}$$

$$\beta = \frac{AP}{IP}$$

$$\theta = \frac{AP}{CP}$$

When these values are substituted in the last equation and AP is cancelled from both sides, we obtain

$$\frac{n_2}{IP} + \frac{n_1}{OP} = \frac{n_2 - n_1}{CP}$$

which is an equation relating the positions of O and I analogous to, but less simple than, the type of equation derived for a spherical mirror in a similar way.

The first conclusion to be drawn from the equation is that a definite (real) image of O is formed at I, since the equation does not contain a term which depends upon the position of A, the point of incidence. It must be stated immediately, however, that in general this is true *only* when A is near to P—in other words, when the rays are paraxial. Rays incident on the refracting surface at a considerable distance from P will not be refracted through I. A ray from O striking the surface at P obviously passes through the surface undeviated.

Another possible case is shown in Fig. 591, where n_2 is again greater

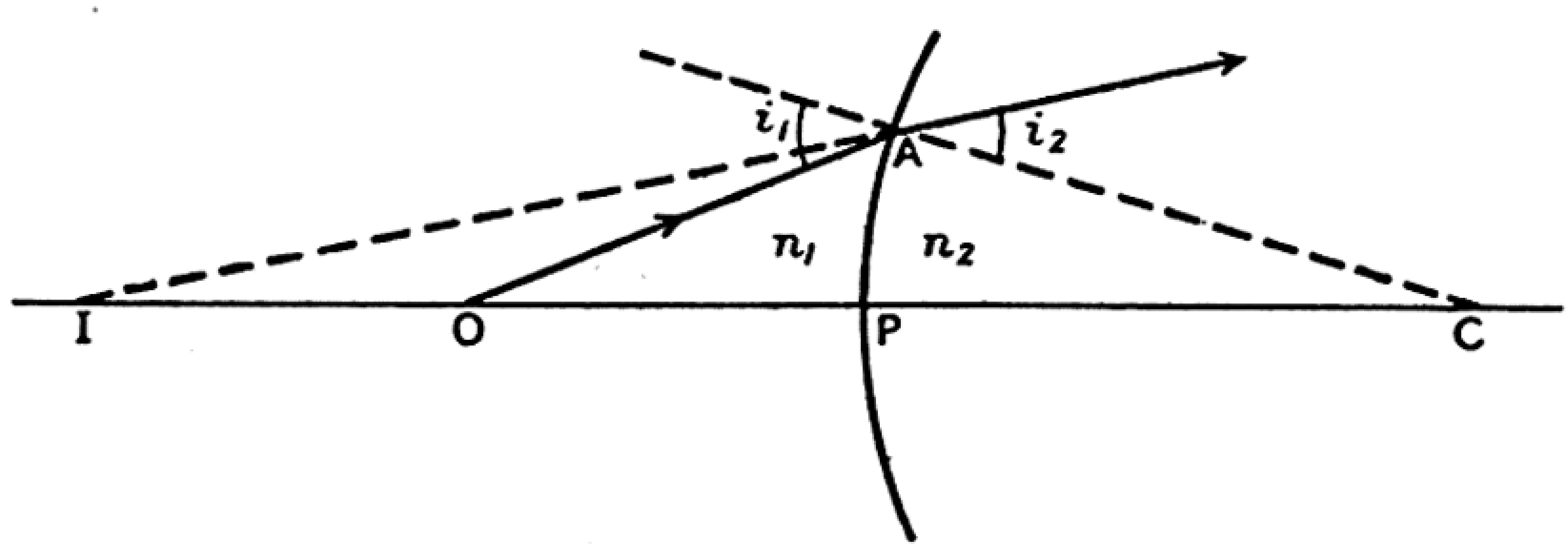


FIG. 591

than n_1 and the surface is again concave towards n_2, but either because the difference of the refractive indices is not so great or because O is nearer to P than before, the refracted ray diverges from the axis so that the image I is virtual. It should be noted, however, that the refracting surface has nevertheless *converged* the light falling on it in the sense that it has reduced the divergence. The student should verify by a method

similar to that just used for the real-image case that Fig. 591 leads to the equation

$$-\frac{n_2}{\mathrm{IP}}+\frac{n_1}{\mathrm{OP}}=\frac{n_2-n_1}{\mathrm{CP}}$$

Thus, as with spherical mirrors, a virtual image is associated with a negative sign in front of the term which represents its position.

The student may also verify that a negative sign appears in front of the object term if the object is virtual. This suggests what in fact is true in all possible cases, namely that as regards the positions of the object and image we can use the "real is positive" sign convention which we adopted for spherical mirrors.

The first two cases we have discussed are illustrated as (i) and (ii) in Fig. 592, and the appropriate equation is written beside the diagram. Fig. 592 (iii) is a case in which the curvature of the surface in relation to the incident light is reversed compared with (i) and (ii), but the relative magnitudes of n_1 and n_2 are also reversed ($n_1 > n_2$) and refraction again results in convergence. It will be realized that, since the path of the light is reversible, O and I can be interchanged, and if this is done in (iii), and n_1 and n_2 are also interchanged, we get (i) both as regards diagram (although this is laterally reversed) and equation.

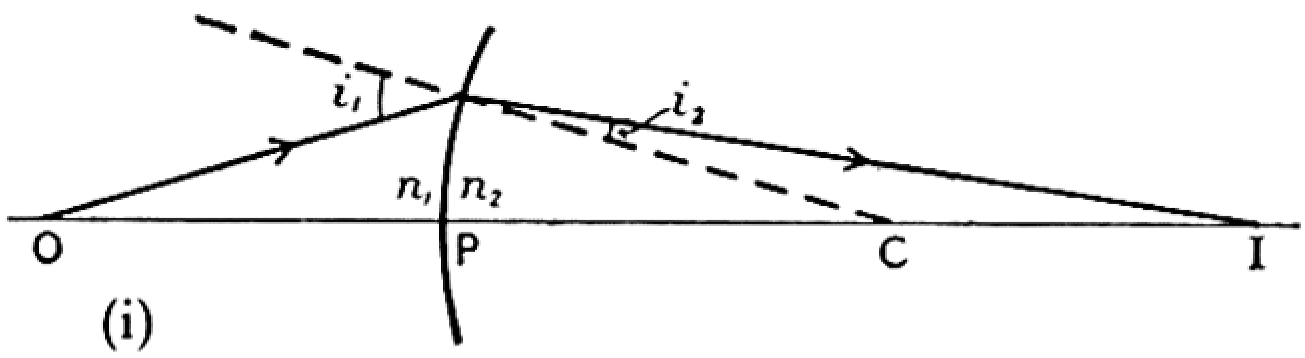
Fig. 592 (iv) and (v) illustrate cases in which refraction produces divergence. The equations attached to each diagram can easily be verified from first principles. If the direction of the light is reversed in any of the cases where I is virtual we obtain the case of a virtual object.

It is now necessary to correlate all the possible cases by means of a sign convention. By writing the equations in the above way we have shown that it is possible to adopt the "real-is-positive" sign convention as regards the positions of the object and image. The right-hand side of the equation is always equal to (difference of refractive indices) ÷ (radius of curvature), and it is necessary to devise a rule which, in conjunction with the R.P. convention, will give this expression the correct sign in every case. We shall adopt the following rule. The radius of curvature CP is always regarded as positive (and will be denoted by r), and the expression must always be written

$$\frac{(\text{refractive index on concave side})-(\text{refractive index on convex side})}{r}$$

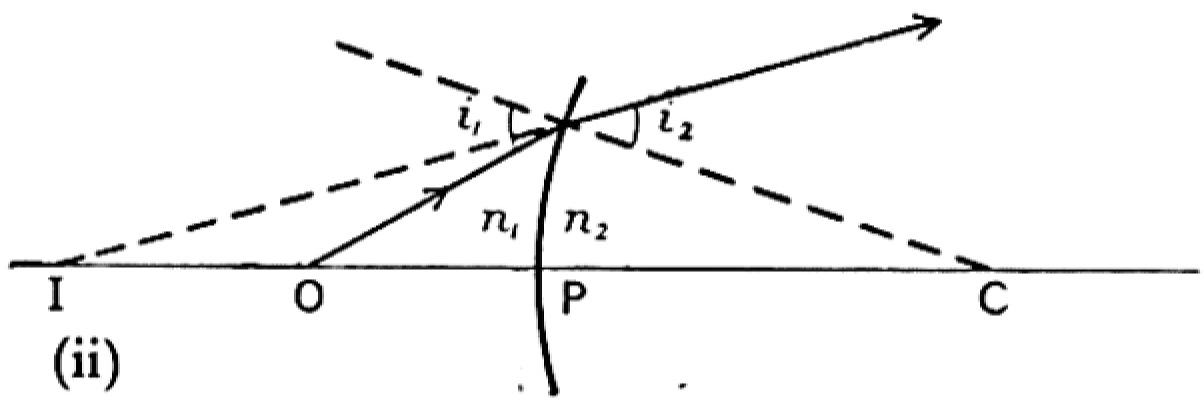
It can easily be verified that this rule does in fact give the correct expression on the right-hand side of the equation in every case. The quantity $\dfrac{(n_{\text{concave}}-n_{\text{convex}})}{r}$ is called the **power** of the refracting surface, and we shall represent it by the symbol K. It will be seen from the equations attached to Fig. 592 that the power is positive when the

I. Refraction produces Convergence



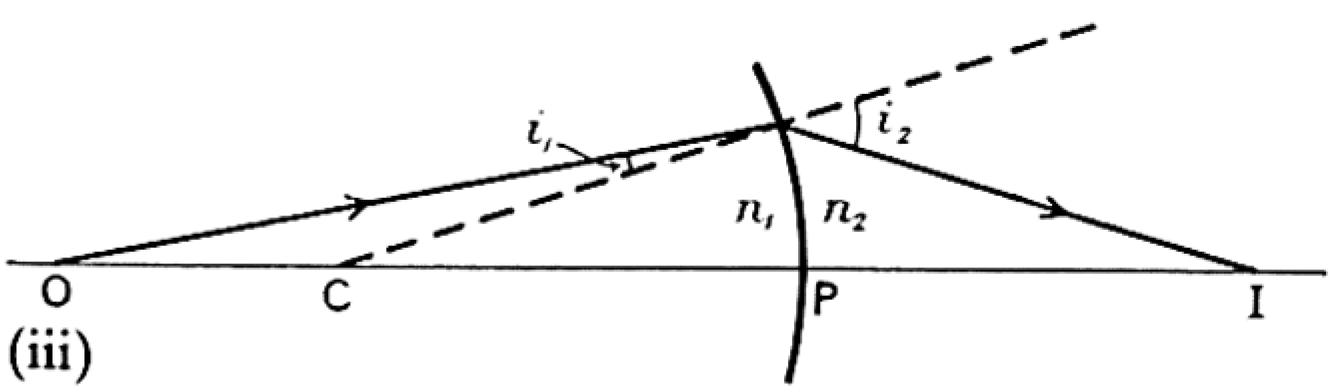
$n_2 > n_1$

$$\frac{n_2}{IP} + \frac{n_1}{OP} = \frac{n_2 - n_1}{CP}$$



$n_2 > n_1$

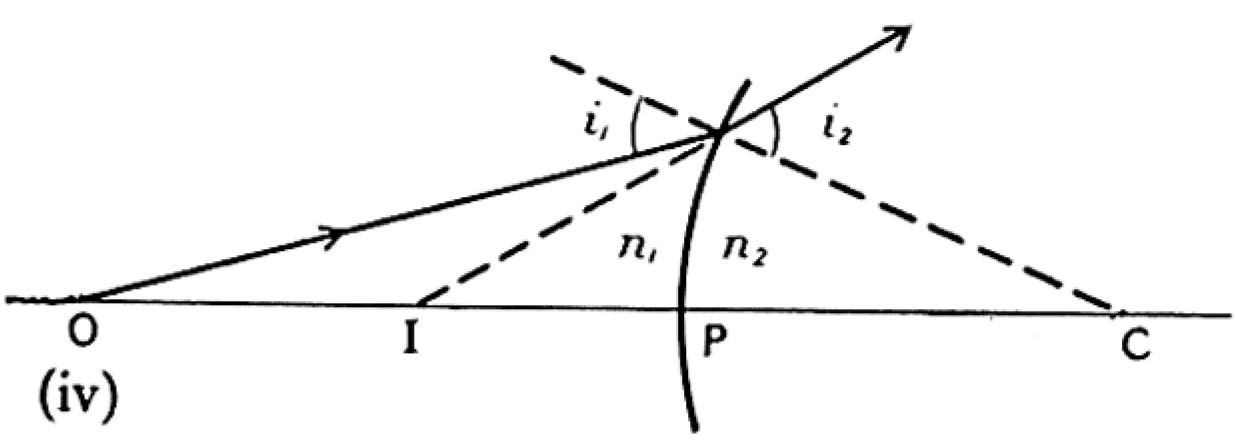
$$-\frac{n_2}{IP} + \frac{n_1}{OP} = \frac{n_2 - n_1}{CP}$$



$n_1 > n_2$

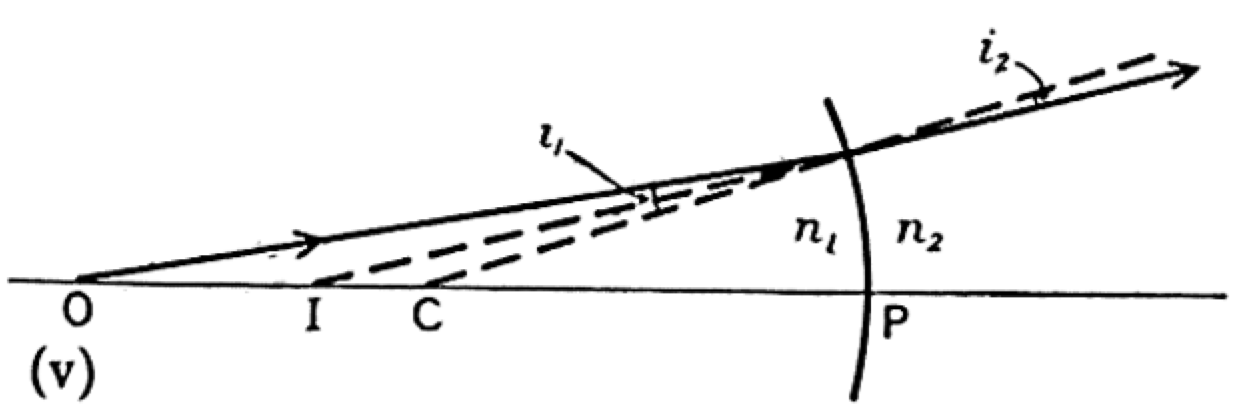
$$\frac{n_2}{IP} + \frac{n_1}{OP} = \frac{n_1 - n_2}{CP}$$

II. Refraction produces Divergence



$n_1 > n_2$

$$-\frac{n_2}{IP} + \frac{n_1}{OP} = \frac{n_2 - n_1}{CP}$$



$n_2 > n_1$

$$-\frac{n_2}{IP} + \frac{n_1}{OP} = \frac{n_1 - n_2}{CP}$$

FIG. 592

surface increases the convergence (or decreases the divergence, which is the same thing) of the incident light, and negative when it increases the divergence (or decreases convergence). Thus we might refer to it as the "converging power" and reckon a negative converging power as a diverging power. The power of a given surface between two given media has the same magnitude and sign irrespective of the direction in which the light passes through it.

If r is expressed in **metres** the power of the surface is expressed in **dioptres.** Thus a spherical surface of radius 20 cm. separating glass (refracting index $=1{\cdot}5$) from air is $\dfrac{1{\cdot}5-1}{0{\cdot}2}$ or $+2{\cdot}5$ dioptres (written $+2{\cdot}5$D) if the glass is on the concave side of the surface and $\dfrac{1-1{\cdot}5}{0{\cdot}2}$ or $-2{\cdot}5$D if the glass is on the convex side.

To summarize, therefore, the general equation applying to all cases of refraction of paraxial rays by a spherical surface separating two transparent media is

$$\frac{\text{(refractive index of medium containing refracted ray)}}{v}+\frac{\text{(refractive index of medium containing incident ray)}}{u}=\frac{\text{(refractive index on concave side of surface)}-\text{(refractive index on convex side of surface)}}{r}$$

where v and u are respectively the distances of the image and object from the refracting surface and are subject to the R.P. sign convention, while r is the radius of curvature of the surface and is always considered positive. We shall write the equation in symbols thus:

$$\frac{n_2}{v}+\frac{n_1}{u}=\frac{n_{\text{concave}}-n_{\text{convex}}}{r}=K \quad . \quad . \quad (1)$$

Principal Foci.—Suppose that, with the light going from medium (1) to medium (2), an object is placed at such a position on the principal axis of a refracting surface that the refracted rays in medium (2) are all parallel to the principal axis. This means that the image is at infinity. We denote this position of the object by F_1, and we call this point the **object focus** or the **first principal focus** for light travelling from medium (1) to medium (2). If we denote the value of u under these conditions (*i.e.* the position of F_1 with respect to the surface) by f_1, the general equation

$$\frac{n_2}{v}+\frac{n_1}{u}=K$$

becomes

$$\frac{n_1}{f_1}=K$$

which gives

$$f_1=\frac{n_1}{K}$$

f_1 is the object focal length or first focal length of the surface for light travelling from (1) to (2).

In the particular case of a *converging* surface f_1 has a *positive* sign, as it should do according to the R.P. convention since it refers to the position of a *real* object. This is illustrated in Fig. 593 (i) where $n_2 > n_1$, so that $K\left(=\frac{n_2-n_1}{r}\right)$ is positive, and $f_1\left(=\frac{n_1 r}{n_2-n_1}\right)$ is also positive. If, as in Fig. 593 (ii), the surface is *diverging* on account of the larger refractive index being on the convex side, f_1 becomes *negative* because the expression for it is then $\frac{n_1 r}{n_1-n_2}$, n_2 being greater than n_1. This occurs

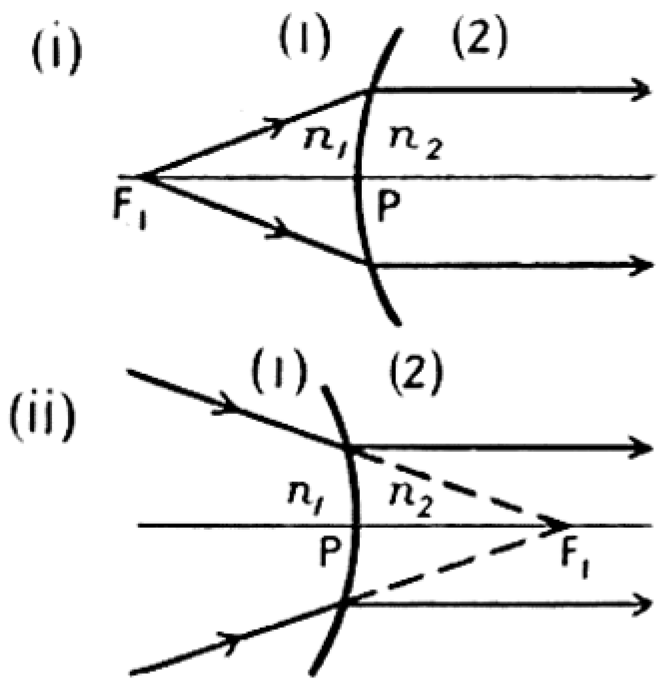


FIG. 593

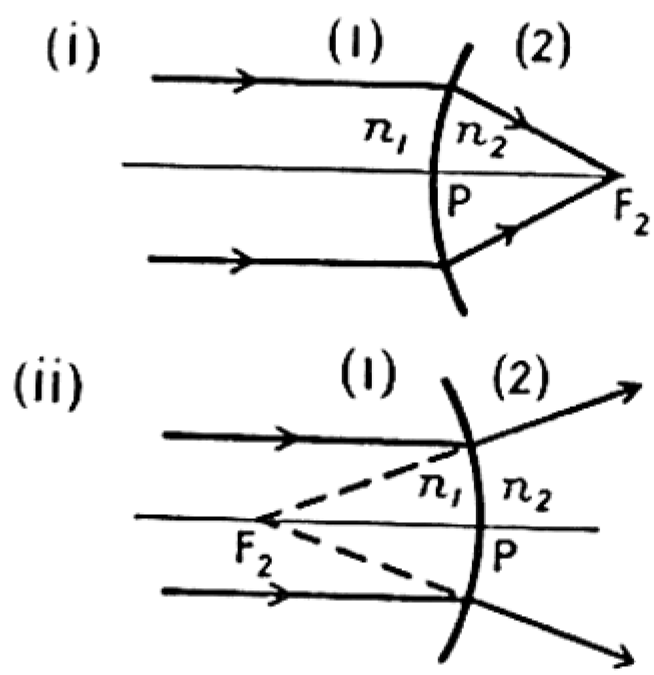


FIG. 594

because, as can be seen from the figure, it is necessary to use a virtual object in order to obtain a parallel beam of light from a diverging surface.

The **second** or **image principal focus** of the surface when light is travelling from (1) to (2) is the position of the image when the incident light is parallel to the principal axis as in Fig. 594 (i) and (ii), in which the position of the second focus is denoted by F_2. If the value of v under these conditions is f_2, we have (remembering that $\frac{1}{u}$ is equal to zero),

$$\frac{n_2}{f_2}=K$$

which gives

$$f_2=\frac{n_2}{K}$$

f_2 is the image focal length or second focal length.

For the converging surface in Fig. 594 (i) in which $n_2 > n_1$ there is a positive second focal length $\left(=\frac{n_2 r}{n_2-n_1}\right)$ as well as a positive first focal length, the positive sign being associated with the converging character

of the surface. In Fig. 594 (ii) the light is again travelling from (1) to (2) and n_2 is again greater than n_1, but the curvature of the surface has been reversed so that medium (1) is now on the concave side and like f_1, f_2 $\left(=\frac{n_2 r}{n_1 - n_2}\right)$ is negative for a diverging surface as we should expect, because it represents the position of a virtual image.

It should be noticed that the object and image focal lengths of a given surface are not equal to each other. However, it can easily be seen that if the direction of the light is reversed in Fig. 593 (i) and (ii), F_1 becomes the image principal focus for light travelling from (2) to (1). Similarly in Fig. 594, F_2 is the object principal focus for light travelling from (2) to (1).

Example.*—An object is situated in liquid (refractive index* 1·3) *on the axis of a spherical refracting surface separating the liquid from glass (refractive index* 1·5). *The distance of the object from the surface is* 39 *cm. and a virtual image is formed* 30 *cm. from the surface. Calculate the radius of curvature of the surface and state which medium is on the concave side of it.*

The ray diagram is shown in Fig. 595, in which the curvature of the surface is not shown since it is not given. It is possible to deduce immediately which way the surface is curved because it evidently produces divergence and therefore has

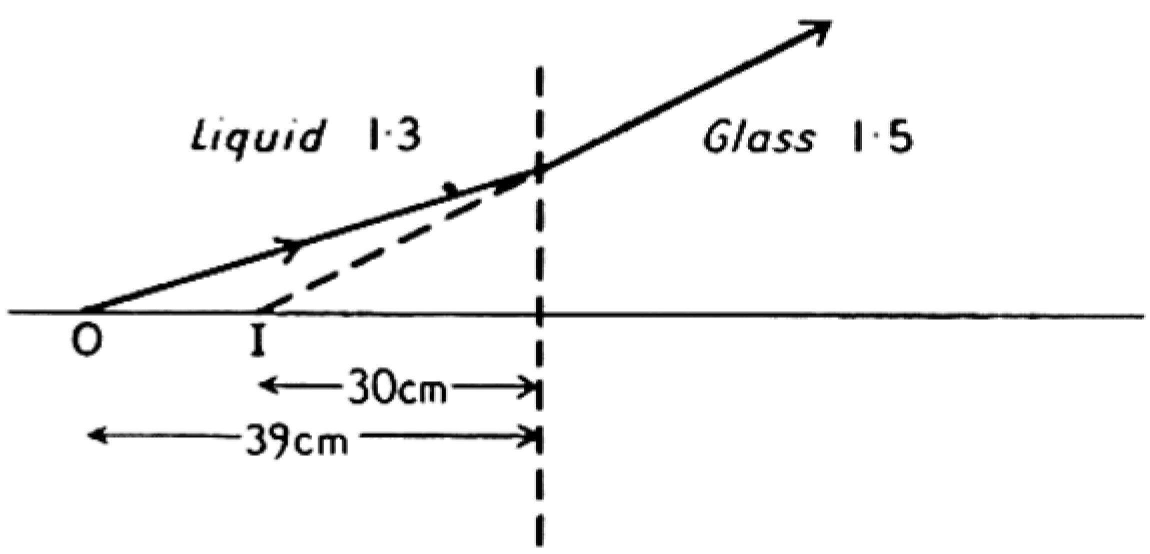


FIG. 595

negative power, which means that the medium with the smaller refractive index (liquid) is on the concave side. We shall see, however, that this fact can also be deduced during the working of the problem.

In equation (1) we have, in this case,

$$n_1 = 1{\cdot}3$$
$$n_2 = 1{\cdot}5$$
$$u = +39 \text{ cm.}$$
$$v = -30 \text{ cm.}$$

Therefore

$$-\frac{1{\cdot}5}{30} + \frac{1{\cdot}3}{39} = \frac{n_{\text{concave}} - n_{\text{convex}}}{r}$$

or

$$-\frac{1}{20} + \frac{1}{30} = \frac{n_{\text{concave}} - n_{\text{convex}}}{r}$$

therefore

$$r = -60(n_{\text{concave}} - n_{\text{convex}})$$

Now according to our rules for signs r is always considered to be a positive quantity, and the last equation shows that, for r to be positive, the difference $(n_{concave} - n_{convex})$ must be negative, which means that the liquid (1·3) must be on the concave side. Thus

$$\begin{aligned} r &= -60(1{\cdot}3 - 1{\cdot}5) \\ &= -60 \times (-0{\cdot}2) \\ &= 12 \text{ cm.} \end{aligned}$$

The radius of curvature is therefore 12 cm. and the liquid is on the concave side.

Refraction by Two or more Surfaces.—If light travelling in medium (1) (Fig. 596 (i)) strikes a curved surface separating medium (1) from medium (2) and forms an image I_1, then I_1 acts as an object for refraction at the surface

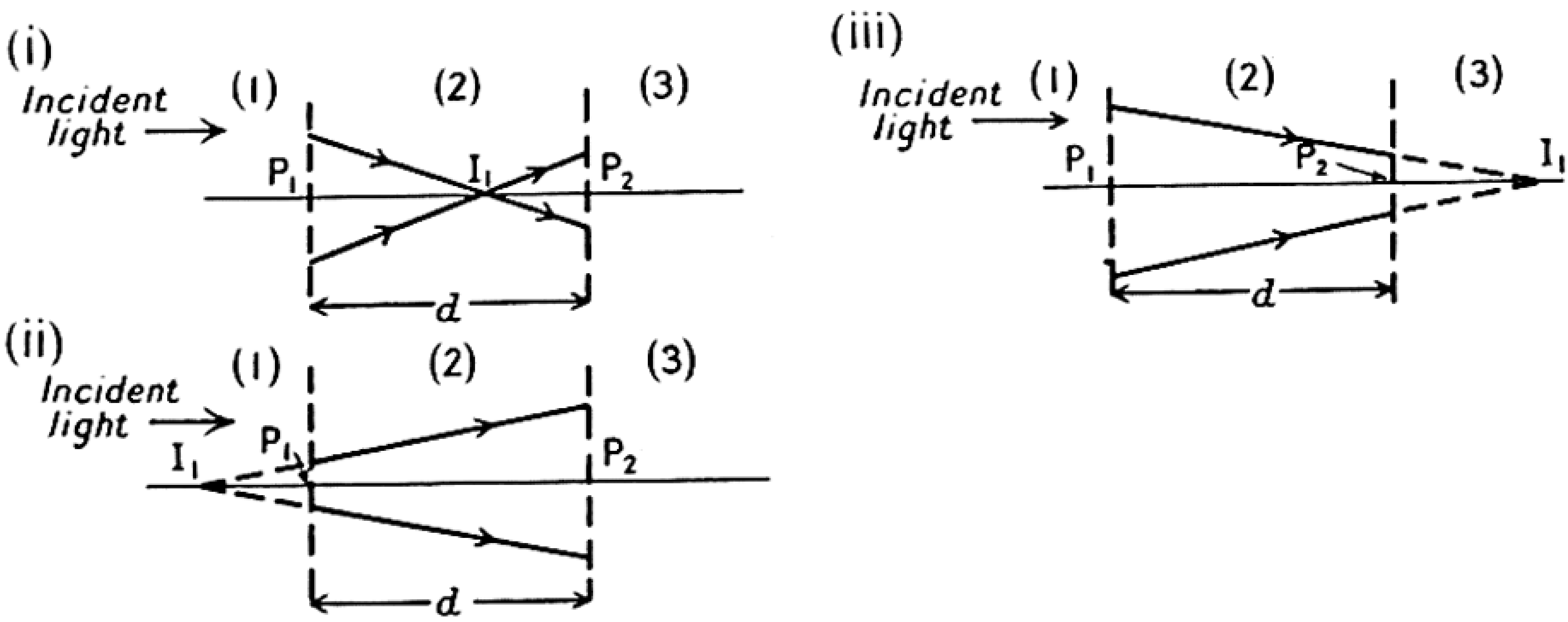


FIG. 596

between (2) and the next medium (3). When (I_1) is situated in medium (2) it is a real image and acts as a real object for the second refraction. Therefore for the first refraction

$$v_1 = +I_1P_1$$

and for the second

$$u_2 = +I_1P_2$$

Consequently the distance between the surfaces d is given by

$$\begin{aligned} d &= I_1P_1 + I_1P_2 \\ &= v_1 + u_2 \end{aligned}$$

In Fig. 596 (ii), I_1 is a virtual image but it acts as a real object for the second refraction. Therefore

$$v_1 = -I_1P_1$$

$$u_2 = +I_1P_2$$

and

$$\begin{aligned} d &= I_1P_2 - I_1P_1 \\ &= v_1 + u_2 \end{aligned}$$

Lastly, in Fig. 596 (iii), I_1 is a real image (or would be were it not for the existence of the second surface) and acts as a virtual object for the second refraction. Therefore

$$v_1 = +I_1P_1$$

$$u_2 = -I_1P_2$$

and

$$d = I_1P_1 - I_1P_2$$

$$= v_1 + u_2$$

Thus the formula $d = v_1 + u_2$ is true for all possible cases when the R.P. sign convention is used. In particular, if the surfaces are very close together so that d is negligible compared with the magnitudes of v_1 and u_2, we have

$$0 = v_1 + u_2$$

so that

$$u_2 = -v_1$$

***Example.**—A point object is situated in air 40 cm. from the surface of a thin-walled hollow glass sphere of diameter 20 cm. filled with water (refractive index $\frac{4}{3}$). Find the position of the image formed by paraxial rays passing right through the sphere.*

In Fig. 597 (i), P_1 and P_2 are the points in which the surface of the sphere intersects the line OC, O being the object and C the centre of the sphere. We can ignore the effect of the glass walls.

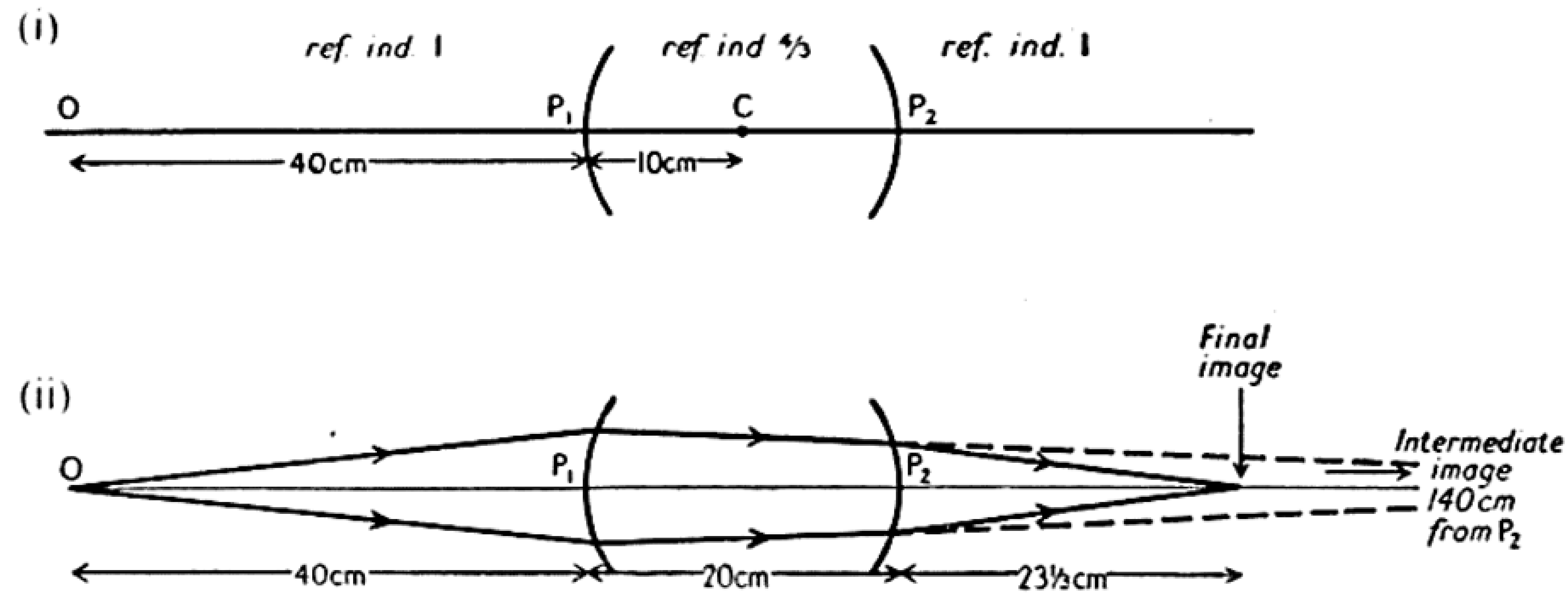


FIG. 597

The first refraction in the neighbourhood of P_1 would form an image whose position with respect to P_1 is v_1. We first find v_1 by substituting in equation (1), thus

$$\frac{\frac{4}{3}}{v_1} + \frac{1}{40} = \frac{\frac{4}{3} - 1}{10}$$

or

$$\frac{\frac{4}{3}}{v_1} = \frac{1}{30} - \frac{1}{40}$$

$$= \frac{1}{120}$$

so that

$$v_1 = 120 \times \tfrac{4}{3}$$
$$= 160 \text{ cm.}$$

Since v_1 is positive the image would be real (were it not for the occurrence of the second refraction) and would be situated 160 cm. to the right of P_1. The light going to this image is, however, intercepted by the second surface in the neighbourhood of P_2, and so the first image constitutes a virtual object for this surface at a distance of $160-P_1P_2$ or 140 cm. from it. For the second refraction, in which water is again on the concave side, we have

$$\frac{1}{v_2} + \frac{\frac{4}{3}}{u_2} = \frac{\frac{4}{3} - 1}{10}$$

where

v_2 is to be found

$u_2 = -140$ cm. since the object is virtual

(It should be noticed that the formula $v_1 + u_2 = d$ (page 782) automatically gives the value of u_2 as $(20 - 160)$, *i.e.* -140 cm.). Therefore

$$\frac{1}{v_2} - \frac{\frac{4}{3}}{140} = \frac{\frac{4}{3} - 1}{10}$$

or

$$\frac{1}{v_2} = \frac{1}{30} + \frac{4}{3 \times 140}$$

so that

$$v_2 = \frac{140}{6}$$
$$= 23\tfrac{1}{3} \text{ cm.}$$

Therefore, since v_2 is positive, the image is real and situated $23\frac{1}{3}$ cm. to the right of P_2. Fig. 597 (ii) is a diagram showing the formation of the intermediate and final images.

Aplanatic Points.—Suppose that the refractive index (n_1) on the concave side of a spherical surface is greater than that (n_2) on the convex

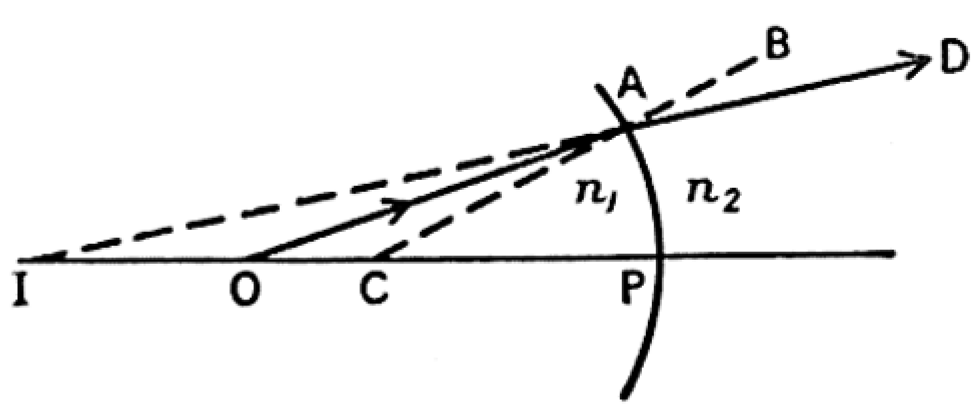


FIG. 598

side, and let an object O be situated on the concave side as in Fig. 598. We consider two points O and I on the principal axis such that

$$OC = \frac{n_2}{n_1} \times CP. \qquad (2)$$

and

$$IC = \frac{n_1}{n_2} \times CP. \qquad (3)$$

and proceed to show that they are object and image respectively and that rays originating from O all diverge from I after refraction irrespective of how far the point of incidence (A) is from P.

Since CP = CA it follows from equations (2) and (3) that

$$\frac{\text{OC}}{\text{CA}}=\frac{n_2}{n_1}=\frac{\text{CA}}{\text{IC}}$$

This relationship between the sides of triangles OCA and ACI together with the fact that the angle $\widehat{\text{OCA}}$ is common to both triangles makes them similar. Therefore, if CA is produced to B and IA to D,

$$\widehat{\text{BAD}}=\widehat{\text{IAC}}=\widehat{\text{AOC}}$$

and

$$\frac{\sin \widehat{\text{OAC}}}{\sin \widehat{\text{BAD}}}=\frac{\sin \widehat{\text{OAC}}}{\sin \widehat{\text{AOC}}}=\frac{\text{OC}}{\text{CA}}=\frac{n_2}{n_1}$$

or

$$n_1 \sin \widehat{\text{OAC}}=n_2 \sin \widehat{\text{BAD}}$$

The angles $\widehat{\text{OAC}}$ and $\widehat{\text{BAD}}$ are therefore related by Snell's law, and, since CA is the normal to the surface at A, they are angles of incidence and refraction respectively for light from O incident at A. The relationship is quite independent of the position of A and is not subject to any approximation or limitation to paraxial rays. All rays diverging from O and striking the surface are refracted so as to appear to come from the image I, no matter how wide the incident pencil from O is. The points O and I which fulfil this condition and whose positions are related by equations (2) and (3) are known as aplanatic points.

2. REFRACTION BY A LENS

General Treatment.—A lens may be said to consist of a piece of transparent material bounded by two curved surfaces through which the light passes in turn. In Fig. 599 the material consists of medium (2) (refractive index n_2). The two curved surfaces are represented by the two dotted lines and are supposed to be spherical, but the direction of the curvature of each is not stated so as to preserve generality. The horizontal line represents the principal axis of the lens, *i.e.* it passes through the two centres of curvature, intersecting the surfaces at P_1 and P_2 respectively. This implies that the curved surfaces are parallel to each other at P_1 and P_2.

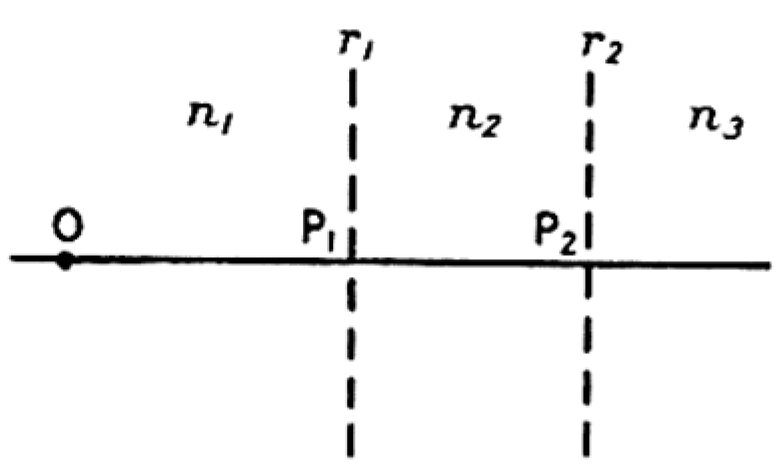


FIG. 599

The distance P_1P_2 is the thickness of the lens, which we shall denote by t, and the radii of curvature of the surfaces are r_1 and r_2 respectively. Let a medium of refractive index n_1 be in contact with one face of the lens and n_3 with the other.

Let paraxial rays from an object O on the common axis arrive at the surface which has a radius r_1 from medium (1). If the position of the object with respect to P_1 is denoted by u_1 and that of the image formed by the first refraction is v_1, we can write

$$\frac{n_2}{v_1}+\frac{n_1}{u_1}=K_1 \qquad (4)$$

This equation, as we know, holds good whether the object is real or virtual, provided that the appropriate sign is inserted when we substitute for u_1. The value of K_1 will be $\frac{n_1-n_2}{r_1}$ if the first refracting surface is concave towards medium (1) and $\frac{n_2-n_1}{r_1}$ if it is concave towards (2). The image whose position with respect to P_1 is v_1 acts as an object for the second refraction and, as we have seen on page 782, the position (u_2) of this object relative to the second surface is given by

$$v_1+u_2=t \qquad (5)$$

provided that the signs of the quantities v_1 and u_2 are governed by the R.P. sign convention and t is always regarded as a positive quantity. Finally, the second refraction takes place according to the equation

$$\frac{n_3}{v_2}+\frac{n_2}{u_2}=K_2 \qquad (6)$$

where v_2 is the position of the final image with respect to P_2, and the value of K_2 is $\frac{n_2-n_3}{r_2}$ if the surface is concave to medium (2) and $\frac{n_3-n_2}{r_2}$ if it is concave to (3).

Equations (4), (5) and (6), taken together, enable any refraction problem of this type to be worked out for paraxial rays.

The Thin Lens.—The term "thin" is applied to a lens when the curved surfaces are so close together that the distance P_1P_2—the thickness of the lens—is negligible compared with the radii of the surfaces. The lens is supposed to be used under conditions in which the distances of the object and final image from it are large compared with its thickness.

The various types of lens are classified according to whether their surfaces are convex or concave towards the medium outside the lens. The possible combinations of concave, convex and plane surfaces together with the name of each lens are shown in Fig. 600. When the thickness

of the material of the lens is greater at the centre than at the edge the lens is **thick-centred,** and it is called a **converging lens** because it causes an increase in the convergence of the light passing through it when the medium on both sides of it is air. Such a lens would, however, produce

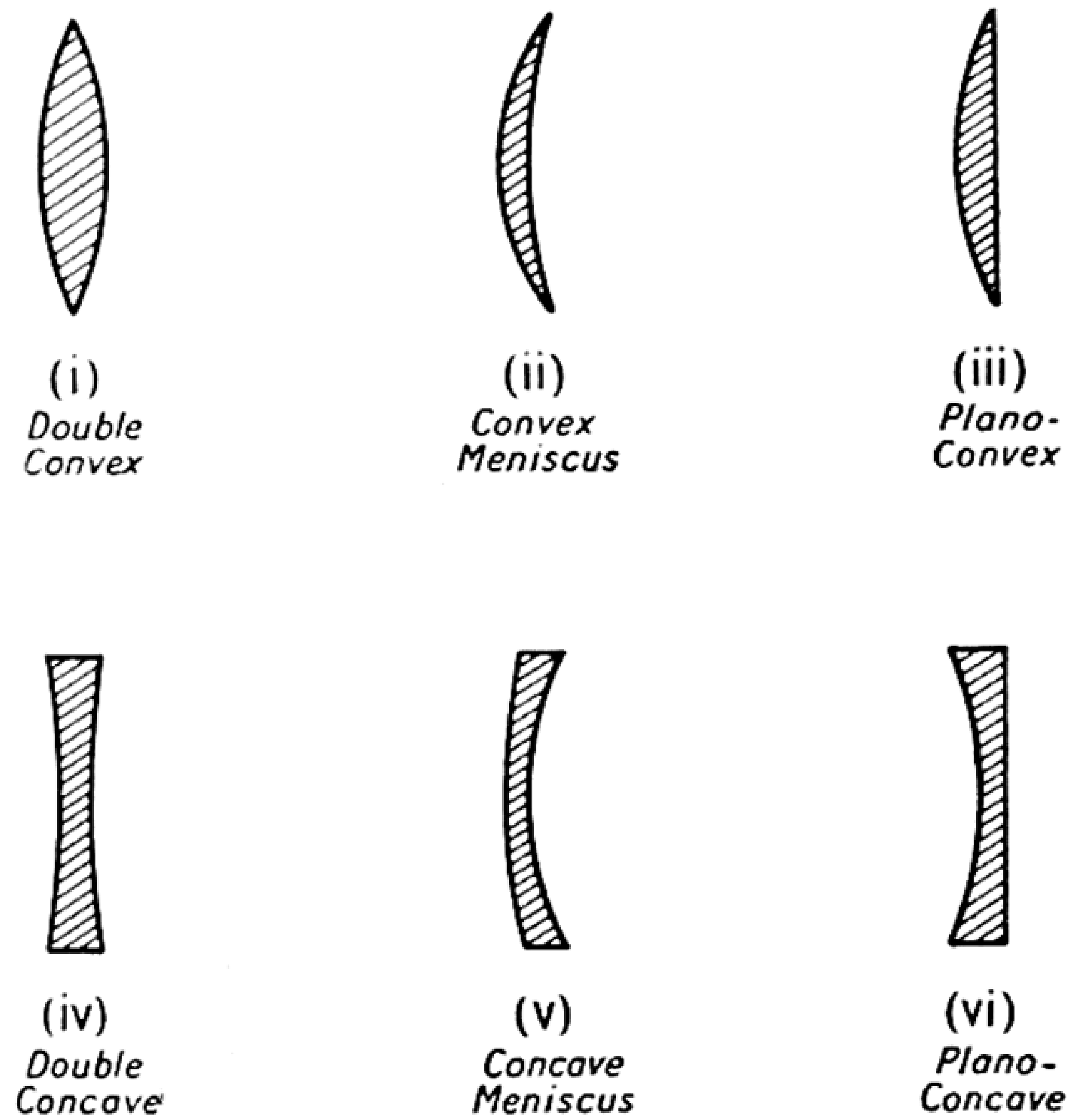


FIG. 600

divergence if it were situated in a medium of greater refractive index than its own material. The opposite type (nos. (iv), (v) and (vi)) are **thin-centred** and **diverging.** The term "meniscus" is applied to lenses which have one concave and one convex surface.

Suppose that light travelling in a medium of refractive index n_1 is incident on a thin lens made of material of refractive index n_2 and, having passed through the lens, enters a third medium n_3.

For the first refraction we can apply equation (4), *i.e.*

$$\frac{n_2}{v_1}+\frac{n_1}{u_1}=K_1 \qquad (7)$$

where u_1 refers to the position of the object relative to the first surface and v_1 to the image formed by that surface. The position u_2 of the object for the second surface is related to v_1 by equation (5), which, since the thickness of the lens t is negligible, becomes

$$v_1+u_2=0$$

or

$$u_2=-v_1$$

Hence the equation for the second refraction (equation (6)) becomes

$$\frac{n_3}{v_2} - \frac{n_2}{v_1} = K_2 \qquad (8)$$

Adding equations (7) and (8) we eliminate v_1 and obtain

$$\frac{n_3}{v_2} + \frac{n_1}{u_1} = K_1 + K_2 \qquad (9)$$

The left-hand side of this equation concerns the positions of the object and final image with respect to the lens and the refractive indices of the media in which the incident and final refracted rays are travelling, but *not* the refractive index of the lens itself. On the right-hand side of equation (9) we have the sum of the powers of the two surfaces. This sum is referred to as the **power of the lens,** and is expressed in dioptres if r_1 and r_2 are expressed in metres. It should be noticed that the power of the lens depends upon the refractive indices of the media on either side of it as well as upon its own, and we should remind ourselves of the expressions for K_1 and K_2 (page 778).

Example.—*A thin lens of glass (refractive index $\frac{3}{2}$) has a convex surface of radius 15 cm. and a concave surface of radius 30 cm. It is placed in an aperture in the vertical wall of a tank containing water (refractive index $\frac{4}{3}$), with its principal axis horizontal and its concave surface towards the water. Find the position of an image of a point object situated on the axis of the lens, in the air, and 20 cm. from the lens.*

The arrangement is shown in Fig. 601 (i). We are given that $n_1 = 1$ (air), $n_2 = \frac{3}{2}$, $n_3 = \frac{4}{3}$, $r_1 = 15$ cm., $r_2 = 30$ cm.

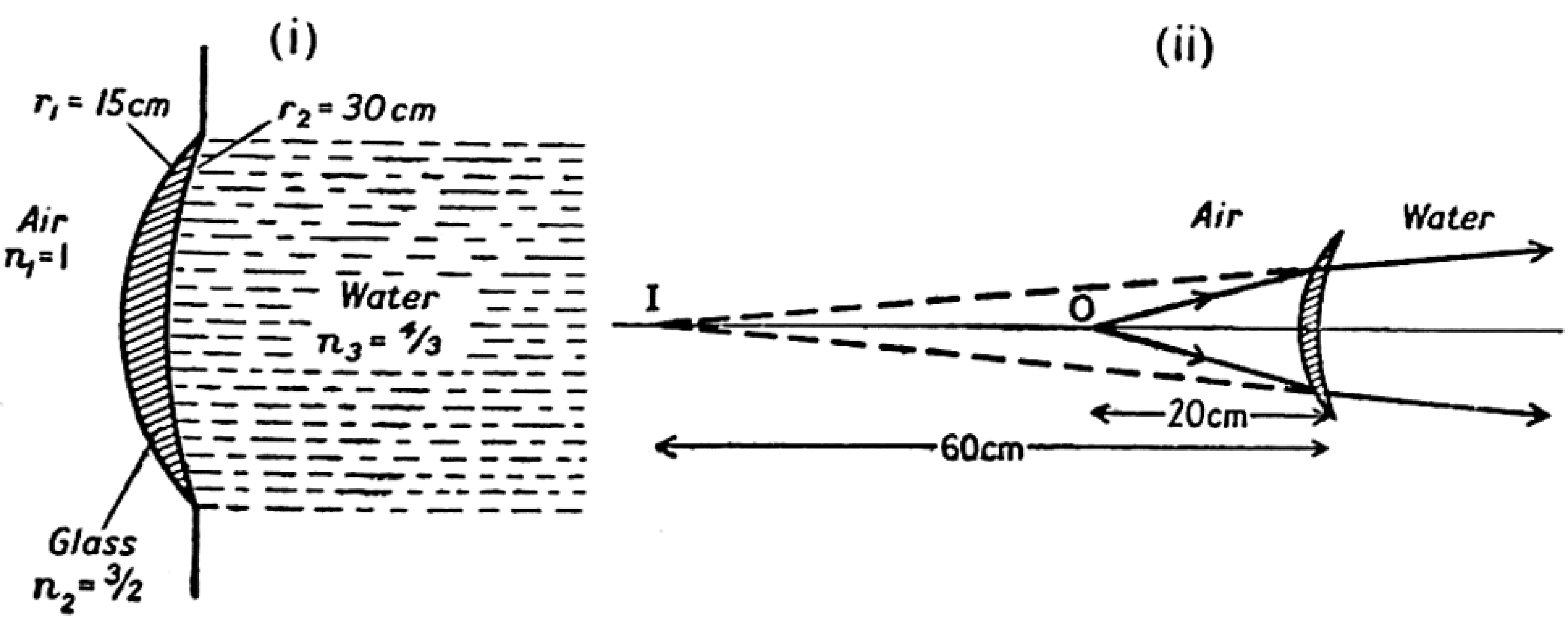


FIG. 601

The power of the air/glass surface K_1 is given by

$$K_1 = \frac{n_2 - n_1}{r_1} = \frac{\frac{3}{2} - 1}{15} = \frac{1}{30}\ \text{cm.}^{-1}$$

since the glass is on the concave side.

The power of the other surface K_2 is given by

$$K_2 = \frac{n_3 - n_2}{r_2} = \frac{\frac{4}{3} - \frac{3}{2}}{30} = -\frac{1}{180} \text{ cm.}^{-1}$$

The total power of the lens is therefore

$$K_1 + K_2 = \frac{1}{30} - \frac{1}{180} = \frac{1}{36} \text{ cm.}^{-1}$$

Since this power is positive, the lens is a converging one when placed in this way between air and water. Applying equation (9) we have

$$\frac{n_3}{v} + \frac{n_1}{u} = \frac{1}{36}$$

where u refers to the position of the object with respect to the lens and is equal to $+20$ cm. since the object is real, and v refers to the image position which is to be found. Substituting in the last equation gives

$$\frac{\frac{4}{3}}{v} + \frac{1}{20} = \frac{1}{36}$$

so that

$$\frac{\frac{4}{3}}{v} = -\frac{1}{45}$$

and

$$\begin{aligned} v &= -\tfrac{4}{3} \times 45 \\ &= -60 \text{ cm.} \end{aligned}$$

The image produced is therefore virtual, and is situated 60 cm. from the lens on the air side. A ray diagram is given in Fig. 601 (ii).

Same Medium on Both Sides of the Lens.—The equation for the case in which the same medium is present on the two sides of the lens can evidently be derived from equation (9). Suppose that the medium surrounding the lens has a refractive index n_1, and that that of the lens material is n_2 as before. The appropriate equation is then derived from equation (9) by substituting n_1 for n_3. We can also drop the subscripts and write u for u_1 and v for v_2 now that we are no longer discussing the two refractions separately. Equation (9) then becomes

$$\frac{n_1}{v} + \frac{n_1}{u} = K_1 + K_2 \qquad . \quad . \quad . \quad . \quad (10)$$

where, of course, u and v are subject to the R.P. convention, and K_1 and K_2 are the powers of the two surfaces of the lens. For a convex surface the power is equal to $\frac{n_2 - n_1}{\text{radius}}$, medium (2) being on the concave side, while for a concave surface the power will be $\frac{n_1 - n_2}{\text{radius}}$. The sum $(K_1 + K_2)$ is the **power of the lens** when it is surrounded by medium (1).

Air on Both Sides of the Lens.—Lenses are, of course, most commonly used with air on each side. Taking the refractive index of air as unity,

we can write down the equation appropriate to this case by putting $n_1 = 1$ in equation (10). We then obtain

$$\frac{1}{v} + \frac{1}{u} = K_1 + K_2$$

where K_1 and K_2 are $\dfrac{n-1}{\text{radius}}$ for a convex surface and $\dfrac{1-n}{\text{radius}}$ for a concave surface. Since n is always greater than unity, the power of a surface which is convex to the air will be positive and that of a concave surface negative.

The power of the lens in air is the sum of the powers of its two surfaces. Thus if a lens has two convex surfaces of radii r_1 and r_2 respectively, its power K will be given by

$$\begin{aligned} K &= K_1 + K_2 \\ &= \frac{n-1}{r_1} + \frac{n-1}{r_2} \\ &= (n-1)\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \quad \text{which is positive.} \end{aligned}$$

If both surfaces have the same radius (r, say), the power of the double convex lens (which is then called equi-convex) is $\dfrac{2(n-1)}{r}$, which is positive. For a pair of concave surfaces of radii r_1 and r_2,

$$\begin{aligned} K &= \frac{1-n}{r_1} + \frac{1-n}{r_2} \\ &= (1-n)\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \quad \text{which is negative.} \end{aligned}$$

For concave surfaces of equal radius r (equi-concave lens), $K = \dfrac{2(1-n)}{r}$, which is negative.

If a lens has a convex surface of radius r_1 and a concave surface of radius r_2, its power is given by

$$\begin{aligned} K &= \frac{n-1}{r_1} + \frac{1-n}{r_2} \\ &= (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \end{aligned}$$

It will be noticed that the power of such a lens would be zero if r_1 were equal to r_2, that is to say it would produce neither convergence nor divergence but would behave like a parallel-sided piece of material.

If the radii of the surfaces are expressed in metres, the powers of the

surfaces and of the lens itself are expressed in dioptres. Thus for an equi-convex lens of glass of refractive index 1·5 with surfaces each of radius 20 cm., each surface has a power of $\frac{1\cdot5-1}{0\cdot2}$ or $+2\cdot5$D, so that the power of the lens is $+5$D. If the surfaces of an equi-concave lens had the same radii, the power of each would be $\frac{1-1\cdot5}{0\cdot2}$ or $-2\cdot5$D and the power of the lens would be -5D.

It must be remembered that u and v must be expressed in the same length unit as the radii. This will be the metre if the powers are expressed in dioptres.

Principal Foci.—For a lens with a power K in air, the object and image positions are related by

$$\frac{1}{v}+\frac{1}{u}=K$$

Suppose that the image is at infinity so that $\frac{1}{v}$ is zero, and that we denote the corresponding position of the object by f_1. The equation then gives

$$f_1=\frac{1}{K}$$

As with the single refraction, so with the lens, we call f_1 the **object** (or **first**) **focal length** of the lens, and we have shown that it is equal to the reciprocal of the power of the lens and it has the same algebraic sign as the power. Fig. 602 (i) shows the position of the first principal focus F_1 or a converging lens for light travelling from left to right. Because the power K is positive in this case, f_1 is positive, as we should expect, since it is necessary to place a real object or its equivalent at F_1 in order that parallel light shall emerge on the other side of the lens. In Fig. 602 (ii) F_1 is the object or first principal focus of a diverging lens for light travelling from left to right. In this case f_1 is negative because the power is negative. The first principal focus is the position of a *virtual* object.

We next consider the **image** (or **second**) **principal focus,** which is the position of the image when the object is at infinity, *i.e.* when a parallel beam is incident on the lens. In this case $\frac{1}{u}$ is zero, and if we denote the corresponding value of v by f_2 we have, from the general equation,

$$f_2=\frac{1}{K}$$

which is the same as the expression for f_1. Therefore for a given lens *in air*, the object and image focal lengths have the same magnitude and sign. Fig. 602 (iii) shows the second principal focus F_2 of a converging lens when light is travelling from left to right. The image is real and f_2 is

positive. For a diverging lens (Fig. 602 (iv)) F_2 is virtual and f_2 is negative. It is evident that for light travelling from right to left in any of the diagrams the positions of the first and second principal foci are interchanged, but for a given lens the position of either is always given by the reciprocal of the power of the lens, the sign of the power denoting whether the focus is real or virtual. From now on, therefore, we shall not usually distinguish between the first and second focal lengths of a lens in air and shall simply refer to *the* focal length, which we shall denote by f. The focal length is positive for a converging lens and negative for a diverging lens, and is always equal to the reciprocal of the power. The lens equation can therefore be written

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

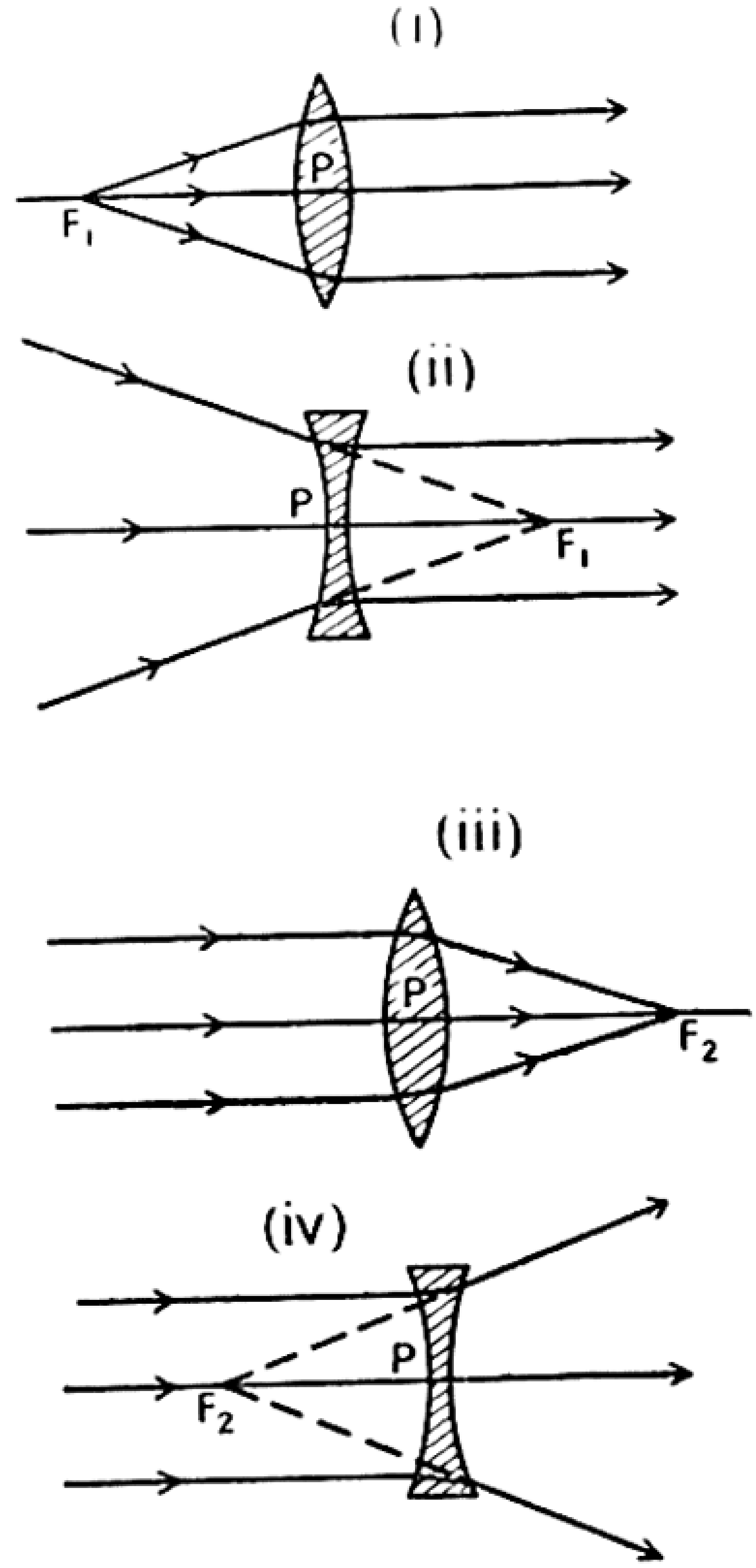


FIG. 602

The student should be able to write out a more direct proof of this equation than we have used—including the expression for f in terms of the refractive index of the material of the lens and its radii of curvature. This is done by establishing the equation for refraction from air to glass at the first surface of the lens (as on page 774) and combining it with the equation for refraction from glass to air at the second surface in a way similar to that on page 786.

Example.*—A thin lens of glass of refractive index* 1·5 *has a convex surface of radius* 50 *cm. A point object on the principal axis* 50 *cm. from the lens forms a virtual image* 25 *cm. from the lens. Find the focal length and power of the lens and the radius and nature of the other surface.*

In the general equation

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}=K$$

we have

$$u=+50 \text{ cm.} \quad \text{(real object)}$$
$$v=-25 \text{ cm.} \quad \text{(virtual image)}$$

Therefore

$$\frac{1}{f}=K=-\frac{1}{25}+\frac{1}{50}$$
$$=-\frac{1}{50}\text{ cm.}^{-1}$$

Therefore f, the focal length, is -50 cm., and K, the power of the lens, is $-\frac{100}{50}$, *i.e.* -2D. The lens is, therefore, diverging.

For the convex surface we have, since the glass is on the concave side of the surface,

$$K_1 = \frac{1{\cdot}5 - 1}{50}$$
$$= +\frac{1}{100} \text{ cm.}^{-1}$$
$$= +1\text{D}$$

If the power of the other surface (in dioptres) is K_2 we have

$$K_2 = K - K_1$$
$$= -2\text{D} - 1\text{D}$$
$$= -3\text{D}$$

Hence the power of the second surface is -3D and, since this is negative, the surface must be diverging and the smaller refractive index (air in this case) must be on the concave side of the surface. Thus the surface is concave to the air, and its radius r is given by

$$-3 = \frac{1 - 1{\cdot}5}{r}$$

where r is in metres. Therefore

$$r = \tfrac{1}{6} \text{ metre}$$
$$= 16\tfrac{2}{3} \text{ cm.}$$

3. FURTHER DISCUSSION OF THIN LENSES

Graphical Construction.—As with spherical mirrors, so with lenses, we can deduce certain rules with regard to the passage of rays through the lens and use them to find the position of the image.

The first rule concerns the properties of the principal foci of the lens. For a converging lens, any ray travelling to the lens from either of the principal foci gives rise to a refracted ray parallel to the principal axis on the other side of the lens (Fig. 603 (i)). Conversely (Fig. 603 (ii)), an incident ray parallel to the principal axis is refracted through the principal focus on the other side of the lens.

For a diverging lens the corresponding rules are shown in Fig. 603 (iii) and (iv). A ray which is travelling towards a principal focus on the opposite side of the lens is refracted parallel to the principal axis, and a ray parallel to the principal axis is refracted so as to appear to come from the focus on the side of the incident light.

These two rules combined are, as we shall see, sufficient to locate the image formed by a given object and lens.

Another rule concerns the passage of a ray through the lens without change of direction. It is evident that a ray travelling along the principal axis fulfils this condition because it strikes each lens surface normally. This is shown on ray (1) in Fig. 604.

We now consider a ray (2) which strikes the lens at the point A_1. Let

A_2 be a point on the second surface of the lens where the surface is parallel to the direction of the first surface at A_1. From this parallelism it follows that the radius C_1A_1 is parallel to C_2A_2 since each is normal to its own surface, and also that (as is shown in the drawing) if the direction of the incident ray at A_1 is such as to cause the refracted ray in the lens to strike the second surface at A_2, then the final emergent ray is parallel to the initial incident ray. The material between A_1 and A_2 behaves as a parallel-sided slab.

Since C_1A_1 is parallel to C_2A_2 it is evident that the triangles C_1A_1C and C_2A_2C are similar, C being the point at which the ray in the lens crosses the principal axis. Therefore we have

$$\frac{C_1C}{C_2C}=\frac{C_1A_1}{C_2A_2}=\frac{C_1P_1}{C_2P_2}$$

This shows that when the incident and emergent rays are parallel, the point C where the ray in the lens crosses the principal axis is a definite point independent of the direction of the incident ray, because the position of C with respect to C_1 and C_2 as given by the last equation is independent of this direction. Thus any other incident ray such as (3) in Fig. 604 which emerges parallel to its original direction must pass through C,

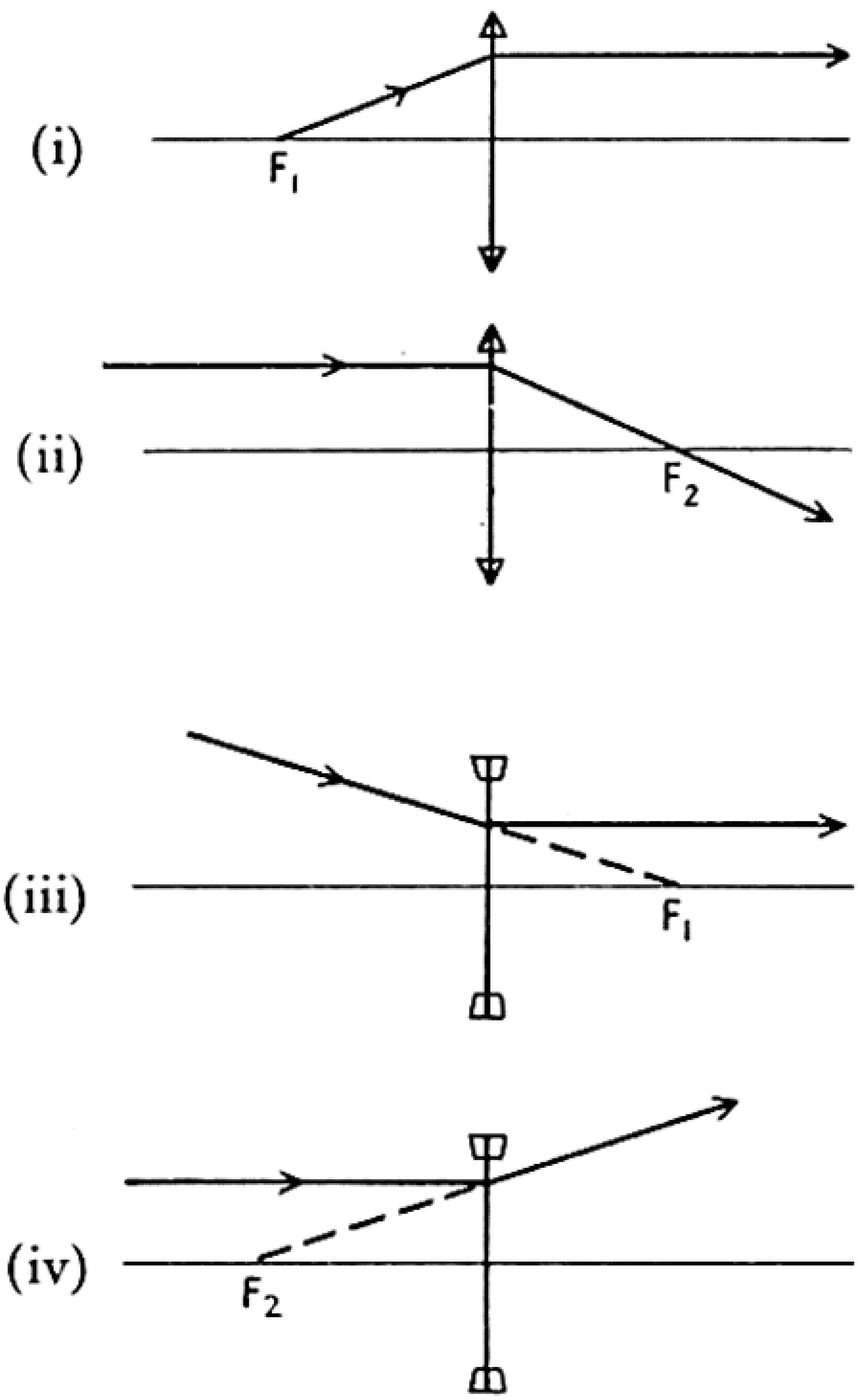


FIG. 603

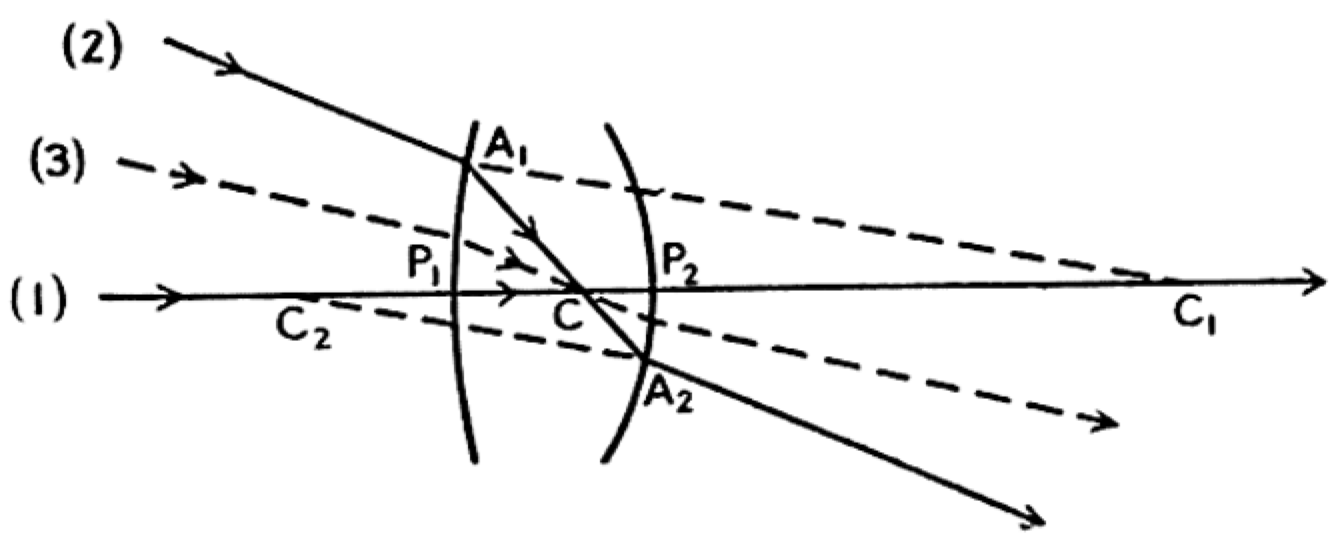


FIG. 604

although the point of incidence on the lens must necessarily be different from that for ray (2). The point C is called the **optical centre** of the lens. The student should find it easy to show that the above equation leads to

the relationship

$$\frac{CP_1}{CP_2}=\frac{C_1P_1}{C_2P_2} \quad . \quad . \quad . \quad . \quad . \quad (11)$$

thus relating the position of C with the points P_1P_2. We may note that when the two surfaces have equal curvatures $C_1P_1=C_2P_2$, so that $CP_1=CP_2$, which means that C is midway between P_1 and P_2.

Equation (11) can be written in the form

$$\frac{CP_1}{(CP_2)+(CP_1)}=\frac{C_1P_1}{(C_1P_1)+(C_2P_2)}$$

or

$$\frac{CP_1}{P_1P_2}=\frac{C_1P_1}{(C_1P_1)+(C_2P_2)}$$

so that

$$CP_1=\frac{(P_1P_2)\times(C_1P_1)}{(C_1P_1)+(C_2P_2)}$$

If we now suppose that the lens is made thinner and thinner so that its thickness P_1P_2 becomes zero, the last equation shows that CP_1 becomes zero, and it can be shown in a similar way—in fact it is obvious—that CP_2 also becomes zero. This means that C coincides with P_1 and P_2 when the lens is indefinitely thin, and this is the state of affairs which we usually suppose to exist when we are dealing with thin lenses. In these circumstances *any* incident ray striking the lens at P_1 is refracted through C and emerges at P_2 without bending or sideways displacement, all three points being coincident. In the case of actual lenses this is sufficiently true for paraxial rays.

The corresponding rule for geometrical construction is therefore as follows. A ray striking the lens at its centre, *i.e.* at the intersection of the lens surfaces with the principal axis, passes through the lens without deviation or displacement.

Fig. 605 (i) and (ii) show the use of the three geometrical construction rules for finding the position of the image I′ of an object O′ situated near, but not on, the principal axis. In (i) the lens is converging and in (ii) it is diverging. The points P_1, P_2 and C are now replaced by the single point P. Let a perpendicular be dropped from O′ on to the axis at O and from I′ to I. We can then show, in the following way, that I is the image of O by proving that their positions are related by the lens equation.

We shall consider Fig. 605 (i). Triangles $OO'F_1$ and PBF_1 are evidently similar, so that

$$\frac{OF_1}{PF_1}=\frac{OO'}{PB} \quad . \quad . \quad . \quad . \quad . \quad (12)$$

But $OF_1 = (OP) - (PF_1)$ and $PB = II'$, so that

$$\frac{OP}{PF_1} - 1 = \frac{OO'}{II'}$$

$$= \frac{OP}{IP}$$

since triangles OO'P and II'P are similar. Dividing through by OP, rearranging the terms and putting PF_1 equal to f, gives

$$\frac{1}{IP} + \frac{1}{OP} = \frac{1}{f}$$

Since this is the form of the lens equation which is appropriate to this case of a converging lens, it follows that I is the image of O. A similar proof, leading to the same result, can be applied to any example of the

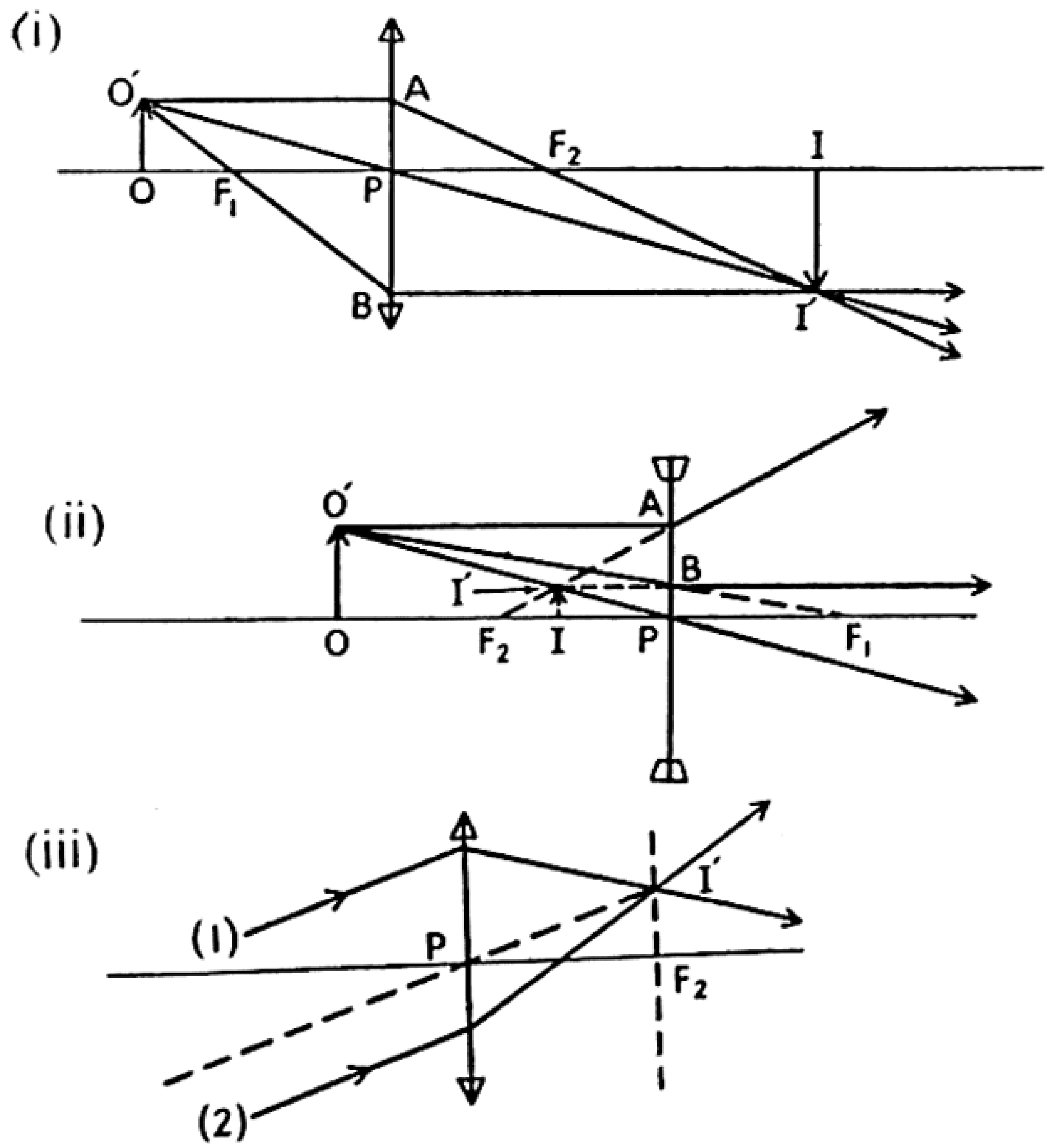


FIG. 605

geometrical construction for both converging and diverging lenses (*e.g.* Fig. 605 (ii)).

In general, therefore, if OO′ is an object lying perpendicular to the principal axis, its image, II′, may be found by using the construction to locate I′, the image of the point O′, and dropping a perpendicular from I′ on to the axis to obtain the position of I.

When a plane object perpendicular to the axis of the lens is situated at a very great distance from the lens its image is formed in the **focal plane,** *i.e.* a plane through the principal focus perpendicular to the axis. Fig. 605 (iii) shows two parallel rays, (1) and (2), from a non-axial point on a distant object being brought to a focus at the image I′ of this point in the focal plane of a converging lens. All incident rays parallel to rays (1) and (2) will be refracted through I′, including the ray through the centre of the lens, which is shown as a dotted line. The fact that this ray is not deviated by the lens provides another useful geometrical construction. Suppose, for example, that it is required to find the direction of *any* incident ray (*e.g.* ray (1)) after passing through the lens. The focal plane is drawn, and a straight line is drawn through the centre of the lens parallel to the incident ray in question. Where this line crosses the focal plane is the point through which the incident ray is refracted by the lens. Thus its direction after refraction can be drawn.

Magnification.—In Fig. 605 (i), triangles OO′P and II′P are similar, so that

$$\frac{II'}{OO'}=\frac{IP}{OP}$$

The ratio $\frac{II'}{OO'}$ is called the **transverse linear magnification,** just as in the case of mirrors (page 725), and the last equation shows that it is equal to the ratio of the distance of the image from the lens to that of the object. The same relationship is true for all other possible cases, *e.g.* Fig. 605 (ii). If, as in the case of spherical mirrors, we express the relation by the equation

$$-m=\frac{v}{u}$$

where v and u are subject to the R.P. sign convention, then m is numerically equal to the ratio of the transverse dimension of the image to that of the object, but it is an algebraic quantity whose sign will be positive when the object and image are the same way up, and negative when one is inverted with respect to the other.

Example.*—A small real object is situated* 40 *cm. from a thin lens on the principal axis. The image is one-third the size of the object. Find the nature and focal length of the lens if the image is* (a) *real,* (b) *virtual.*

(*a*) Since the object and image are both real there will be inversion, and

$$m=-\frac{1}{3}$$

Therefore

$$\frac{1}{3}=\frac{v}{u}$$

and since

$$u = +40 \quad \text{(real object)}$$

$$\therefore \quad v = +\frac{40}{3} \text{ cm.}$$

It should be noted that v automatically acquires its correct sign when the above-mentioned convention as to the sign of m is adopted.

Substituting in the lens equation we obtain

$$\frac{3}{40} + \frac{1}{40} = \frac{1}{f}$$

whence

$$f = +10 \text{ cm.}$$

Therefore the lens is converging and has a focal length of 10 cm.

(*b*) In this case there is no inversion, so that

$$m = +\frac{1}{3}$$

which gives, since $u = +40$ cm.,

$$v = -\frac{40}{3} \text{ cm.}$$

Therefore

$$-\frac{3}{40} + \frac{1}{40} = \frac{1}{f}$$

so that

$$f = -20 \text{ cm.}$$

Therefore the lens is diverging and its focal length is -20 cm.

It is instructive to solve simple lens problems by graphical construction. Fig. 606 (*a*) illustrates the method in (*a*) of the above example. The arrow

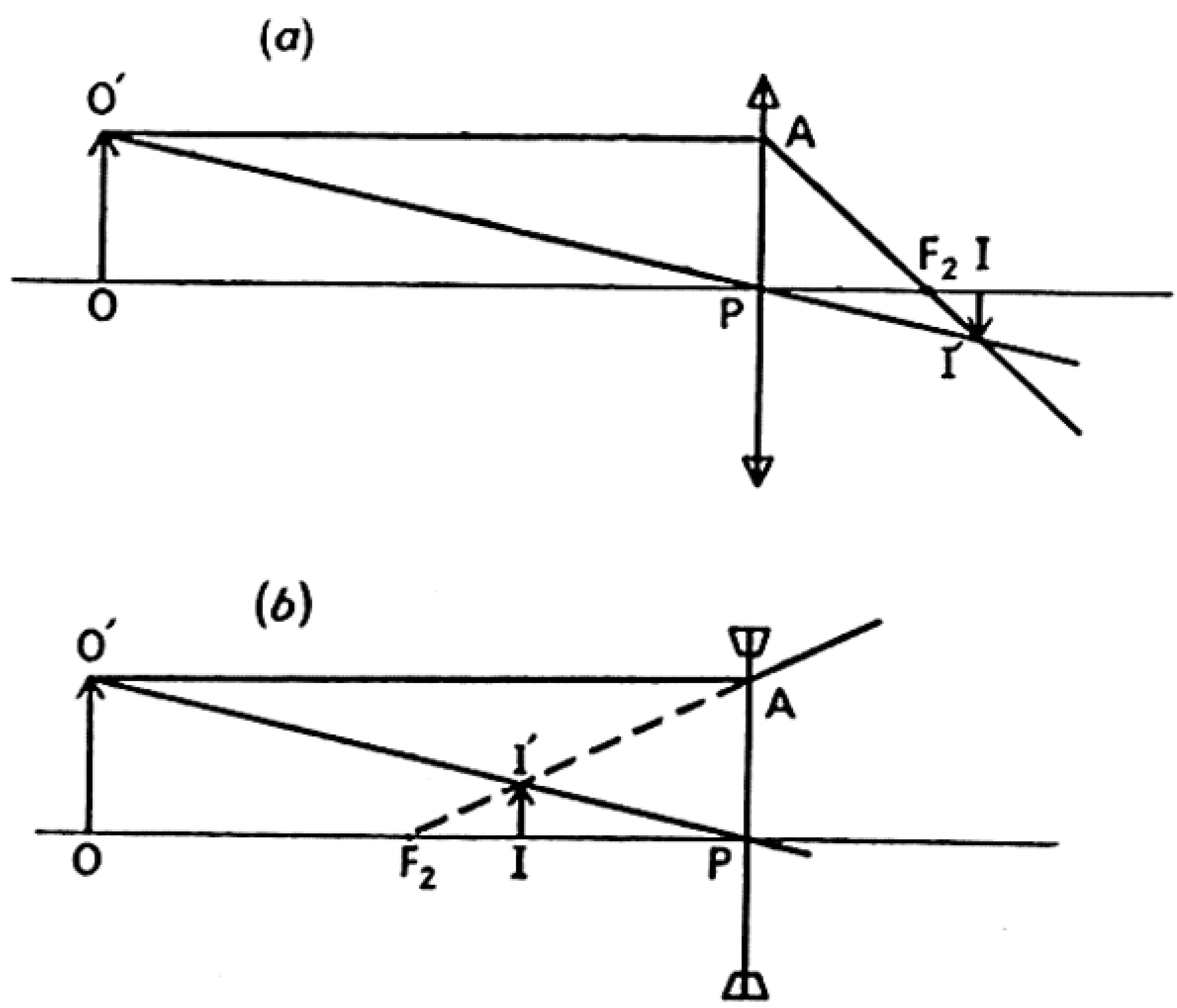


FIG. 606

OO′ of any convenient height is drawn at a distance of 40 cm. from the lens, and the image is drawn on the other side of the lens with its head I′ on

O′P produced and its tail I on the principal axis at such a distance from P as to make II′ one-third of the height of OO′. This implies, of course, that IP must be equal to $\frac{40}{3}$ cm., and this fact may be used in drawing II′. Next, O′A is drawn parallel to OP and A is joined to I′. The line AI′ will cross the axis at the focus F_2, and PF_2 is the required focal length.

The geometrical method would be similar for (*b*) and is illustrated in Fig. 606 (*b*).

Nature and Position of Image for Various Positions of the Object.— For a converging lens it is easily seen from the lens equation or from the geometrical construction that a real inverted image is formed by a real object situated on the principal axis anywhere *except* between the two principal foci, *i.e.* for any value of u greater than f. As the real object is brought in from infinity to the object (first) principal focus, the real image moves out from the image (second) principal focus to infinity on the other side of the lens. The image is diminished (*i.e.* there is a magnification of less than unity) for all positions of the object between infinity and $2f$ from the lens. At this last position (as can be verified by the equation) the image is the same size as the object (v also being equal to $2f$). When u is reduced from $2f$ to f, v increases from $2f$ to infinity and the image is always magnified. The object and image are conjugate

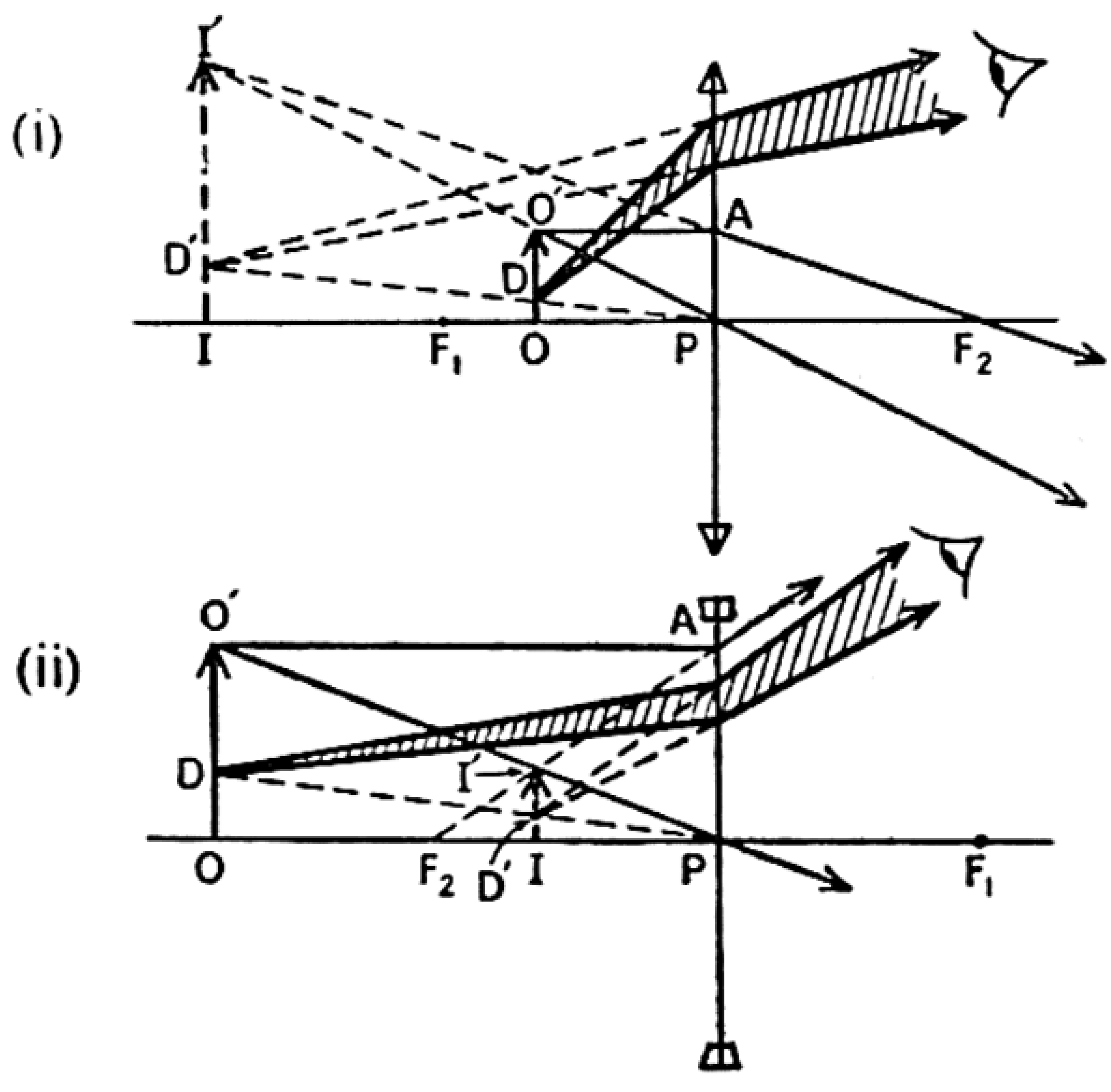


FIG. 607

foci, their positions being interchangeable owing to the fact that the path of any ray through the lens can be reversed. When the object is between the principal focus of a converging lens and the lens itself (u less than f), the image is virtual, erect and magnified (Fig. 607 (i)) and, of course, on the same side of the lens as the object. The image can be seen only by

looking through the lens. A cone of rays is shown in the drawing by which an eye sees a point D′ which is the image of D. A real object point and its virtual image are conjugate foci, becoming respectively a real image and a virtual object when the light is reversed. The student should make a ray diagram of the latter case.

A real object always gives rise to a diminished erect virtual image with a diverging lens (Fig. 607 (ii)), and the image always lies between the image focus and the lens, since the top of it must always lie on the line F_2A. Reversal of the directions of the rays changes the image into a virtual object and the object into a real image.

Newton's Formula.—Consider the case of the formation of a real image I of a real object O by a converging lens as shown in Fig. 608 (i). The distance of O from the object principal focus is OF_1 and is given by

$$OF_1 = OP - PF_1$$

and since u is equal to $+OP$ and f is equal to $+PF_1$ we can write this relationship as

$$x_1 = u - f$$

where x_1 is numerically equal to the distance OF_1 but has an algebraic sign determined by this last equation. In Fig. 608 (ii), in which a diverging lens gives rise to a virtual image of a real object, the object and image foci are (as explained on page 790) on opposite sides of the lens as compared with case (i), and the distance of O from its appropriate principal focus is given by

$$OF_1 = OP + PF_1$$

which can again be written

$$x_1 = u - f$$

because

$$u = +OP \quad \text{and} \quad f = -PF_1$$

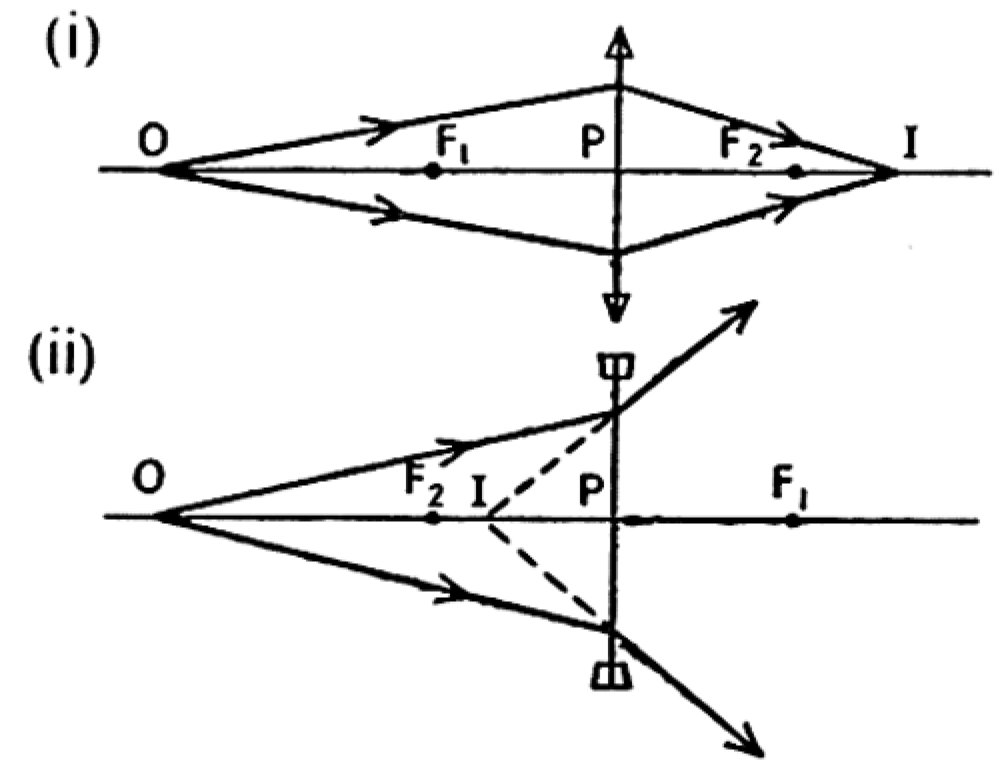


Fig. 608

By considering all possible cases, including those in which O is placed on the side of F_1 opposite to the side from which the light is coming and those in which O is a virtual object (u negative), we conclude that in the equation $x_1 = u - f$, x_1 is always numerically equal to the distance of O from F_1 but that, as a result of the application of the real-is-positive sign convention to u and f (*not* directly to x_1 itself), x_1 is negative when O is on the side of F_1 *to* which the light is travelling, *i.e.* on the right-hand side of F_1 in Fig. 608 (i) and (ii).

The distance of the image from the image focus F_2 is I_2F_2, and we can write

$$x_2 = v - f$$

where x_2 is numerically equal to IF_2 and has an algebraic sign determined by this last equation. By examining the variation of the sign of $(v-f)$ with the magnitude and sign of v and f, it will be found that x_2 is negative when the image is situated on the side of F_2 *from* which the incident light is coming.

The dependence of the signs of x_1 and x_2 with the positions of the object and image is shown in the diagram in Fig. 609 (i) which refers to a converging lens and in Fig. 609 (ii) for a diverging lens.

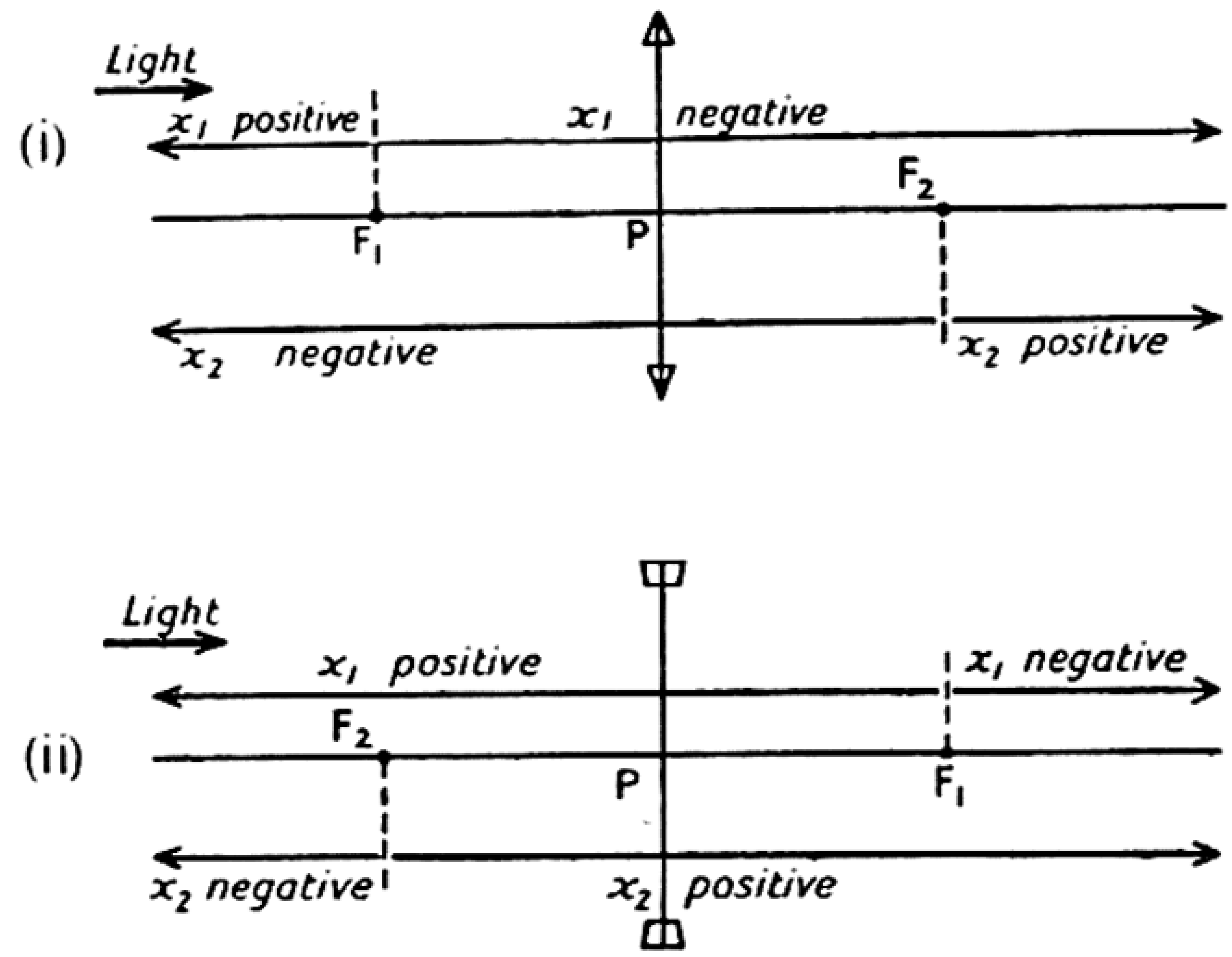


FIG. 609

The reason for introducing the focal distances x_1 and x_2 is to establish a very neat formula due to Newton. This is done as follows. We have the three equations

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

$$x_1=u-f$$

$$x_2=v-f$$

Substituting for u and v in the first equation we obtain

$$\frac{1}{x_2+f}+\frac{1}{x_1+f}=\frac{1}{f}$$

This gives

$$f(x_1+x_2+2f)=(x_1+f)(x_2+f)$$

which evidently gives

$$x_1x_2=f^2$$

This is Newton's formula, and it is particularly useful in connection with lenses which cannot be regarded as thin. This is so because x_1 and x_2

refer to the positions of object and image relative to the principal foci and not to the lens itself.

Whether the lens is converging or diverging f^2 is positive, which means that x_1 and x_2 must always be either both positive or both negative. The meaning of this fact in terms of the possible simultaneous positions of the object and image can be realized by inspection of Fig. 609. If the object lies to the left of F_1 the image must lie on the right of F_2 (x_1 and x_2 positive), and if the object is on the right of F_1 the image is on the left of F_2 (x_1 and x_2 negative).

Other Formulæ for Magnification.—Combining the equations

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

and

$$-m=\frac{v}{u}$$

we obtain

$$1-m=\frac{v}{f}$$

$$\therefore\quad -m=\frac{v-f}{f}$$

$$=\frac{x_2}{f}$$

In a similar way we have

$$-\frac{1}{m}+1=\frac{u}{f}$$

which leads to

$$-m=\frac{f}{x_1}$$

It is evident that Newton's formula can be derived by equating these two expressions for m.

***Example* (i).**—*A diverging lens of focal length* 10 *cm. produces an image having a magnification of* 5. *Find the position of the image and the nature and position of the object if the image is* (a) *real,* (b) *virtual.*

(*a*) The combination of the lens equation with $-m=\frac{v}{u}$ gives, as we have just seen,

$$1-m=\frac{v}{f}$$

In this case $f=-10$ cm., and m is ± 5 according to whether the object is virtual (no inversion) or real (inversion). Thus, substituting in the above equation, we have

$$v=-10(1\mp 5)$$
$$=10(-1\pm 5)$$

and since the image is given as real, v must be positive, and this decides that the positive sign must be taken, otherwise v would be negative. Therefore

$$v = 10(-1+5)$$
$$= +40 \text{ cm.}$$

The image is therefore 40 cm. from the lens on the side away from the incident light (since it is real). Furthermore,

$$\frac{v}{u} = -m$$

so that

$$\frac{40}{u} = -5$$
$$\therefore\ u = -8 \text{ cm.}$$

Therefore the object is virtual and is situated 8 cm. on the side of the lens away from the incident light.

(*b*) Again we have

$$v = 10(-1 \pm 5)$$

and since v is negative this time, we must take the negative sign. Thus

$$v = 10(-1-5)$$
$$= -60 \text{ cm.}$$

Therefore the virtual image is 60 cm. from the lens on the same side as the incident light. Also

$$\frac{v}{u} = -m$$

so that

$$-\frac{60}{u} = +5$$
$$\therefore\ u = -12 \text{ cm.}$$

which means that the object is virtual and lies 12 cm. on the side of the lens away from the incident light.

***Example* (*ii*).**—*A lens forms a virtual image of a real object. When the object is moved so as to move the image 30 cm. further from the lens, the numerical value of the magnification is increased by 2. Find the focal length and nature of the lens.*

Suppose that in the first position the image is p times the size of the object and is situated d cm. from the lens. Then, since the image is virtual, there is no inversion and

$$m_1 = +p$$

and

$$v_1 = -d$$

because the image is virtual. But

$$1 - m_1 = \frac{v_1}{f}$$

$$\therefore\ 1 - p = -\frac{d}{f} \qquad \text{(1)}$$

When the image is moved 30 cm. away from the lens, its distance from the lens becomes $(d+30)$ cm. and its size becomes $(p+2)$ times that of the object.

Therefore

$$m_2 = p + 2$$

and

$$v_2 = -(d+30)$$

Substituting these values in the equation

$$1 - m_2 = \frac{v_2}{f}$$

we obtain

$$1 - (p+2) = -\frac{(d+30)}{f} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Subtracting equation (2) from equation (1) gives

$$2 = \frac{30}{f}$$

$$\therefore f = +15 \text{ cm.}$$

The lens has a focal length of 15 cm. and is converging.

This example emphasizes what is implicit in the equation used, namely that the magnification changes by unity when the image moves through a distance equal to the focal length of the lens.

Deviation Due to a Lens.—Consider a ray parallel to the axis of a lens striking the lens at a point B distant y from P (Fig. 610 (i) and (ii)), where y is small compared with the focal length of the lens.

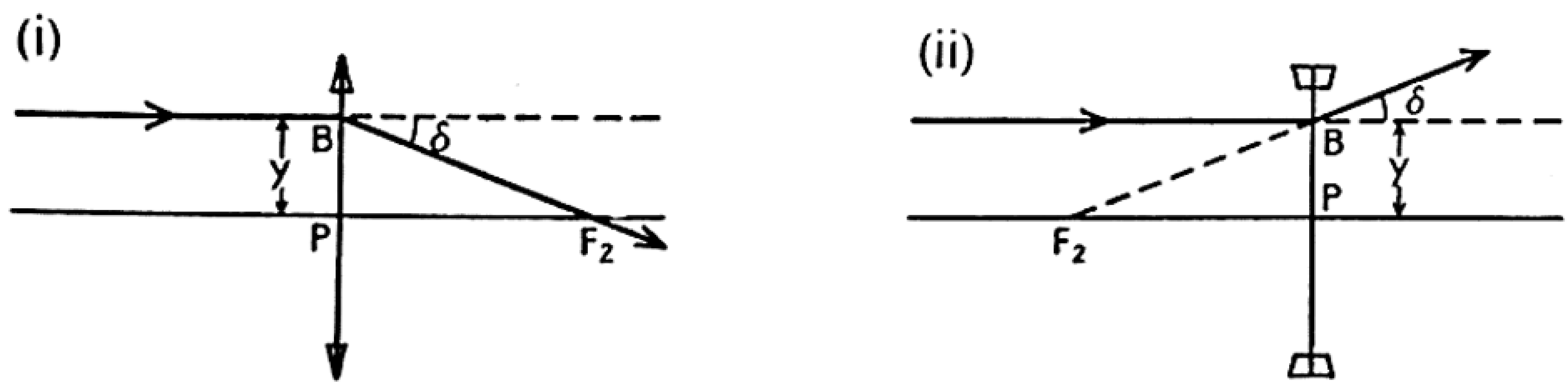


FIG. 610

In the case of a converging lens (Fig. 610 (i)) the refracted ray passes through the second principal focus F_2 and the deviation which the ray experiences is the angle δ. Evidently

$$\delta = B\widehat{F_2}P$$

$$= \frac{y}{PF_2} \quad \text{since } y \text{ is small}$$

Inspection of Fig. 610 (ii) for a diverging lens shows that the same relationship is true in this case.

If we write

$$\delta = \frac{y}{f}$$

where y is always considered positive and f is subject to the R.P. convention, then a deviation *towards* the axis (which happens with a converging

lens) is positive and a deviation away from the axis (diverging lens) is negative.

The material of the lens at any point such as B can be regarded as a "thin" prism (page 756), the refracting angle (α) of which is the angle between the tangents to the two faces of the lens at B (Fig. 611). The deviation δ produced by such a prism is shown on page 756 to be given by

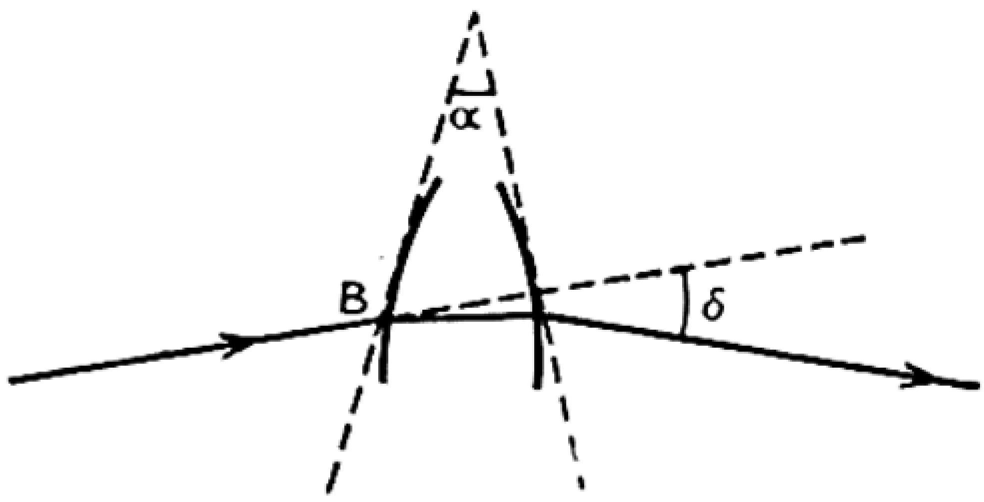


FIG. 611

$$=(n-1)\alpha$$

and is independent of the angle of incidence on the first face provided it is small. This means that the deviation produced by a given lens when the ray is incident at a given distance from P is the same for all positions of the object provided paraxial conditions exist. Three positions are shown in Fig. 612 for a converging lens, the angle δ being the same in each case and equal to $\frac{BP}{PF_2}$ from (ii). Corresponding drawings can be made for a diverging lens.

In Fig. 612 (i) we have

$$\delta = B\widehat{O}P + B\widehat{I}P$$

$$= \frac{BP}{OP} + \frac{BP}{IP} \quad \text{if BP is small}$$

But

$$\delta = \frac{BP}{PF_2}$$

so that

$$\frac{1}{OP} + \frac{1}{IP} = \frac{1}{PF_2}$$

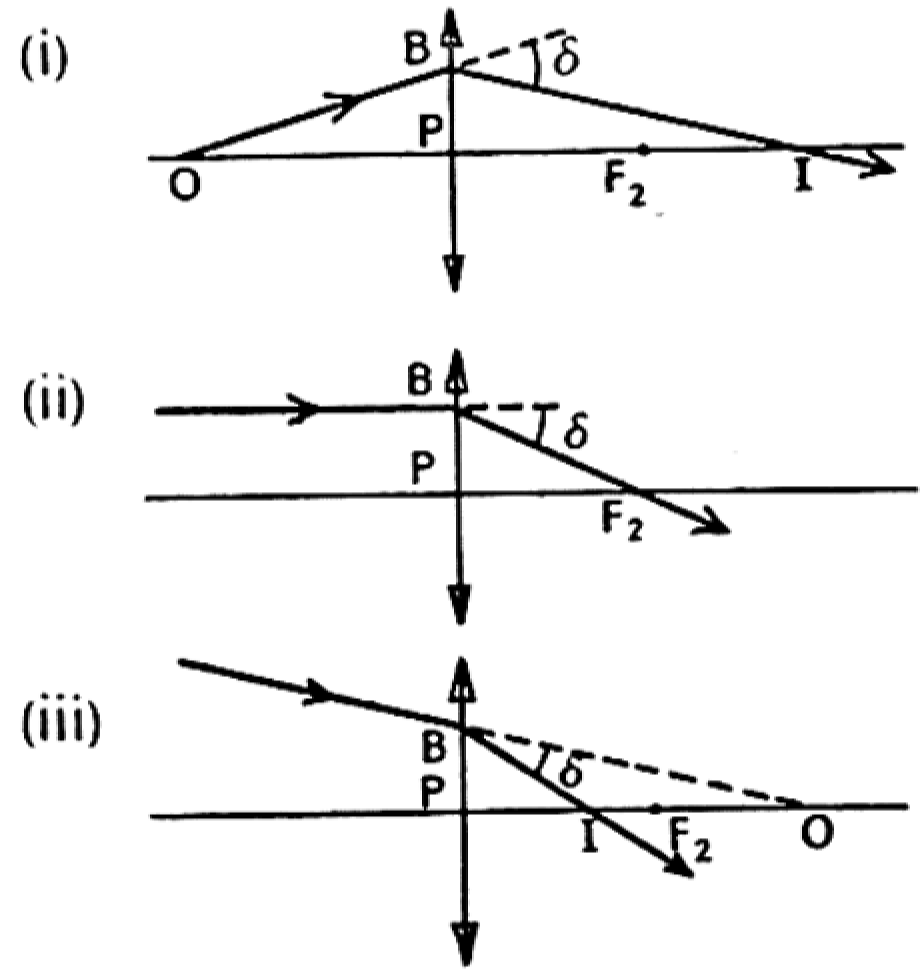


FIG. 612

This is the lens equation, but the foregoing is *not* an adequate proof of it unless it is shown that, at any rate for paraxial rays, PF_2 is independent of the position of B. The actual value of α at any point B can be derived by geometry in terms of BP and the radii of the lens surfaces, and hence the formula for the focal power in terms of the curvatures of the surfaces and the refractive index can be established. If this is done it is found that, for spherical lens surfaces, as B moves away from P, the deviation increases (from zero at P) in such a way as to cause all paraxial rays from a given point

O on the axis to pass through (or diverge from) a single point I on the axis. In particular all rays parallel and near to the axis are brought to a common focus. When paraxial conditions are not satisfied spherical aberration occurs (page 815).

Combination of Thin Lenses in Contact.—Suppose that two thin lenses, of focal lengths f' and f'', are placed in contact, their principal axes being coincident as in Fig. 613. In order to preserve generality we shall work with the symbols u, v and f rather than consider a particular case. Let u represent the position of an object on the axis relative to the lens on which the light is incident first, and suppose that this is the lens of focal length f'. Let the position of the image which would have been formed in the absence of the second lens be represented by v'. Then adopting the real-is-positive convention we have

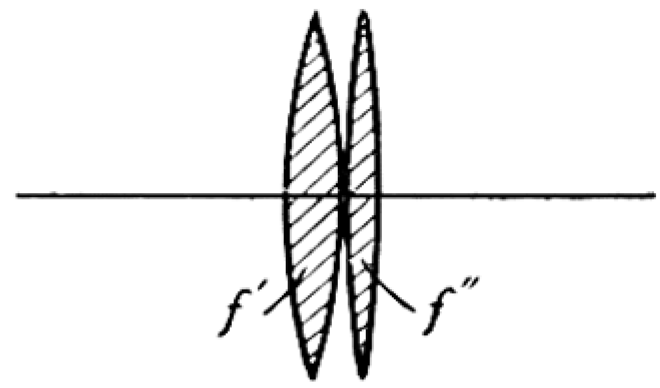


FIG. 613

$$\frac{1}{v'}+\frac{1}{u}=\frac{1}{f'} \quad . \quad . \quad . \quad . \quad . \quad (13)$$

The first image acts as an object for the second lens, and if the position of this object is represented by u' we have, as on page 786,

$$u' = -v'$$

since the distance between lenses is zero. This relationship is evidently true because the numerical values of u' and v' are equal and their signs must be opposite, because if the first image is real it acts as a virtual object for the second lens and if the image is virtual it acts as a real object. Therefore in place of the equation

$$\frac{1}{v}+\frac{1}{u'}=\frac{1}{f''}$$

for the second lens (v representing the position of the final image) we can write

$$\frac{1}{v}-\frac{1}{v'}=\frac{1}{f''} \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Adding equations (13) and (14) we obtain

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f'}+\frac{1}{f''}$$

Therefore the combination behaves like a single lens whose focal length f is given by

$$\frac{1}{f}=\frac{1}{f'}+\frac{1}{f''}$$

This is a very important formula, relating the equivalent focal length of a combination of thin lenses with the focal lengths of the separate lenses. It can, of course, be extended to more than two lenses, except for the fact that the thickness of the combination may then cease to be negligible as we have supposed it was in the foregoing. The formula is true regardless of the algebraic signs of f' and f'' provided these are correctly inserted.

It can be seen in this connection how useful the idea of the power of a lens is. The formula simply represents the fact that the power of a combination of thin lenses in contact is equal to the algebraic sum of the powers of the separate lenses. Thus a converging lens of power $+5\text{D}$ ($f' = +20$ cm.) combined with a diverging lens of power -3D ($f'' = -33\frac{1}{3}$ cm.) gives a combination of power $(5-3)\text{D}$, *i.e.* $+2\text{D}$ ($f = +50$ cm.), which is converging.

* **Two Thin Lenses not in Contact. Thick Lens.**—Evidently two separated thin lenses through which light passes successively can be treated individually, the image formed by the first lens acting as the object for the second, and many elementary problems can be tackled in this way. The combination can, however, be regarded as a single refracting unit having an effective focal length of its own. When its properties are discussed from this point of view it is known as a "thick lens," because it behaves like a single lens in which the distance between the surfaces is not negligible. We shall only indicate briefly the method of approach.

In Fig. 614, the two separated coaxial lenses are taken to be converging lenses—it is necessary to make a choice of the type of lens in order to draw

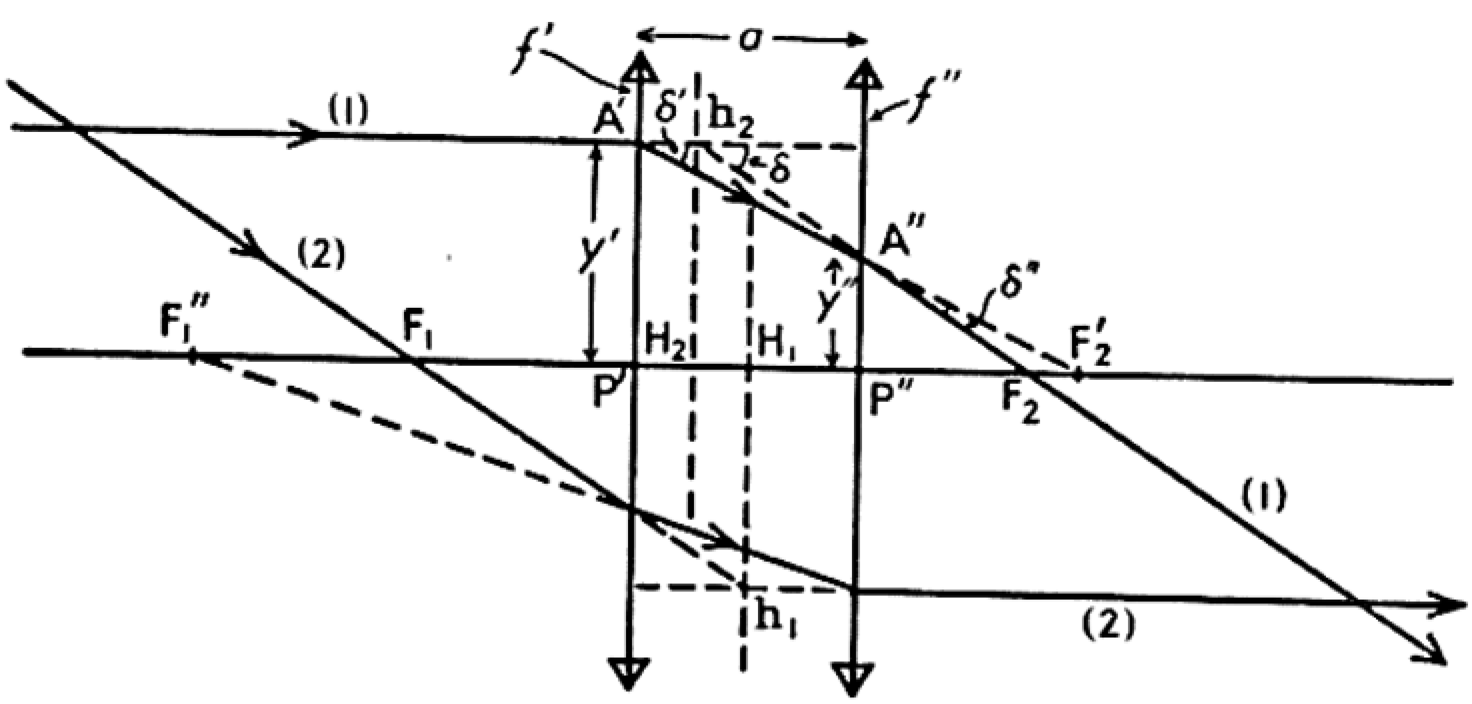


Fig. 614

a diagram—but the following discussion can be adapted to both possible types of lens, and the formulæ are universally applicable when the sign convention is used. A ray (labelled (1) in the diagram) travels parallel to the common axis of the lenses, striking the first lens (focal length f') at a *small* distance y' from the centre of the lens. It is deviated through the angle δ', and in the absence of the second lens it would then pass

* This section is somewhat more advanced than the general standard of this book.

through the image principal focus F_2' of the first lens. The deviation is given by

$$\delta' = \frac{y'}{f'}$$

The ray strikes the second lens (focal length f'') at a distance y'' from the axis. The deviation produced (δ'') is given by

$$\delta'' = \frac{y''}{f''}$$

and the ray emerging from the second lens crosses the axis at F_2 which is, by definition and by analogy with the thin lens, the image focus of the combination, *i.e.* it is the position of the image of an infinitely distant object. The total deviation δ produced by the combination is obtained by continuing the incident ray forward and the emergent ray backward (as indicated by the dotted lines) to meet at the point h_2. The line h_2H_2 is drawn perpendicular to the axis, and it evidently represents the plane at which the total deviation would take place if it occurred in a single step. This is called the second **principal plane** of the combination. If this deviation is regarded as being due to a single lens which replaces the combination, then H_2 is the position of the centre of this lens, the second principal focus of which is F_2, so that its focal length is H_2F_2, which we shall denote by f. It must be understood however that the replacement of the combination by a single thin lens situated at H_2 is only valid for a ray of the type of ray (1) or its direct reverse.

A second incident ray (2) is in such a direction that it emerges from the combination parallel to the axis, so that F_1, where the incident ray crosses the principal axis, is the first or object focus of the combination. The line h_1H_1 represents the first principal plane of the combination, and H_1 is the position of the centre of a thin lens by which the combination could be replaced as regards rays of the type of ray (2).

We now proceed to derive an expression for the focal length f of the thin lens situated at H_2, which would be equivalent to the combination as regards rays of the type of ray (1).

Evidently the total deviation δ of ray (1) is given by

$$\begin{aligned}\delta &= \delta' + \delta'' \\ &= \frac{y'}{f'} + \frac{y''}{f''}\end{aligned}$$

But for the thin lens situated at H_2 the deviation is given by

$$\delta = \frac{y'}{f}$$

where f is numerically equal to H_2F_2. Thus

$$\delta = \frac{y'}{f} = \frac{y'}{f'} + \frac{y''}{f''}$$

or

$$\frac{1}{f} = \frac{1}{f'} + \frac{y''}{y'} \cdot \frac{1}{f''}$$

Clearly the triangles $A'P'F_2'$ and $A''P''F_2'$ are similar, so that

$$\begin{aligned} \frac{y''}{y'} &= \frac{P''F_2'}{P'F_2'} \\ &= \frac{P'F_2' - P'P''}{P'F_2'} \\ &= \frac{f' - a}{f'} \end{aligned}$$

where a is the distance between the lenses. Therefore

$$\begin{aligned} \frac{1}{f} &= \frac{1}{f'} + \frac{f' - a}{f'} \cdot \frac{1}{f''} \\ &= \frac{1}{f'} + \frac{1}{f''} - \frac{a}{f'f''} \quad . \quad . \quad . \quad . \quad . \quad (15) \end{aligned}$$

In the special case we have considered, f' and f'' are both positive. They would have negative values for diverging lenses. The separation of the lenses a is always positive.

Suppose that we proceeded in the above way to find an expression for the focal length of the single thin lens which, placed at H_1, would be equivalent to the combination as regards ray (2). We should then arrive at exactly the same result because, if ray (2) is reversed, the principle is exactly the same as for ray (1) except that f' and f'' would be interchanged, and inspection of equation (15) shows that this would not alter the equation.

We conclude, therefore, that the combination has an effective focal length (f) given by equation (15) irrespective of which of the two lenses the incident light strikes first, but the first and second foci are situated at a distance f from the first and second *principal planes* respectively and not from, say, the two outer surfaces of the combination.

In general, too (unless $f' = f''$), the two foci are not symmetrically placed with regard to the separate lenses, which means that the principal planes are not symmetrical because the foci *are* symmetrical with respect to these planes. It is quite possible, depending upon the values of f', f'' and a, for the principal planes to have any relative position and for one or both of them to be situated outside the combination. In Fig. 614

the principal planes are "crossed" because the incident light traverses the second plane first.

As a simple particular case of a lens combination, suppose that the power of the combination $\left(\frac{1}{f}\right)$ is zero. This means that incident rays parallel to the axis will emerge parallel to the axis, having been neither diverged nor converged. The combination is called a **telescopic system,** and the separation of the lenses a is then given in terms of the focal lengths by

$$0=\frac{1}{f'}+\frac{1}{f''}-\frac{a}{f'f''}$$

or

$$a=f'+f''$$

which means that two of the principal foci are coincident, the separation of the lenses being the algebraic sum of the focal lengths. For a pair of converging lenses the arrangement and action of the telescopic combination is as shown in Fig. 615 (i), in which the second focus F_2' of the

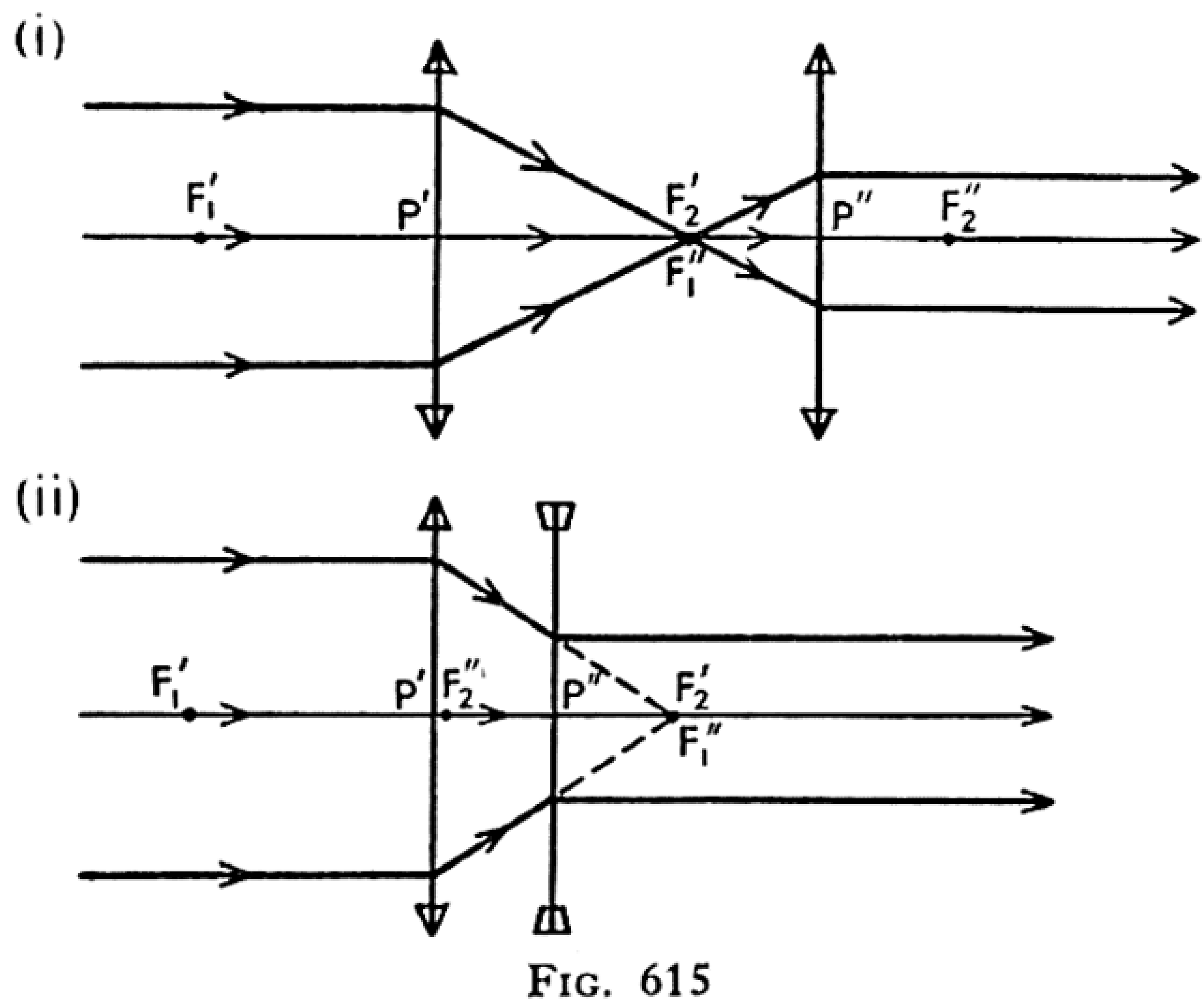


FIG. 615

first lens and the first focus F_1'' of the second lens coincide, $P'F_2'$ being numerically equal to f' and $P''F_1''$ to f''. In Fig. 615 (ii) a similar effect is obtained with a combination of a converging lens with a diverging lens of numerically smaller focal length, the common focus being at F_2', F_1''. Since a cannot be negative the focal length of the diverging lens is necessarily the shorter and, for the same reason, it is impossible to make a telescopic combination with two diverging lenses.

Two rays of the same types as (1) and (2) in Fig. 614 are shown in Fig. 616 (i). The principal planes are represented by the dotted vertical

lines passing through the principal points H_1 and H_2, and the positions of the lenses are not shown. For simplicity the principal planes are not shown as crossed as they are in Fig. 614. The rays are drawn *as though* the total deviation of each occurred at the appropriate principal plane. Evidently any such pair of rays can be used to show the formation of an image by the combination, the point object O′ being the point from which the rays originate and the image I′ being the point at which they cross after passing through both lenses. If perpendiculars are dropped on to the axis, OO′ is a finite (but small) object and II′ is its image.

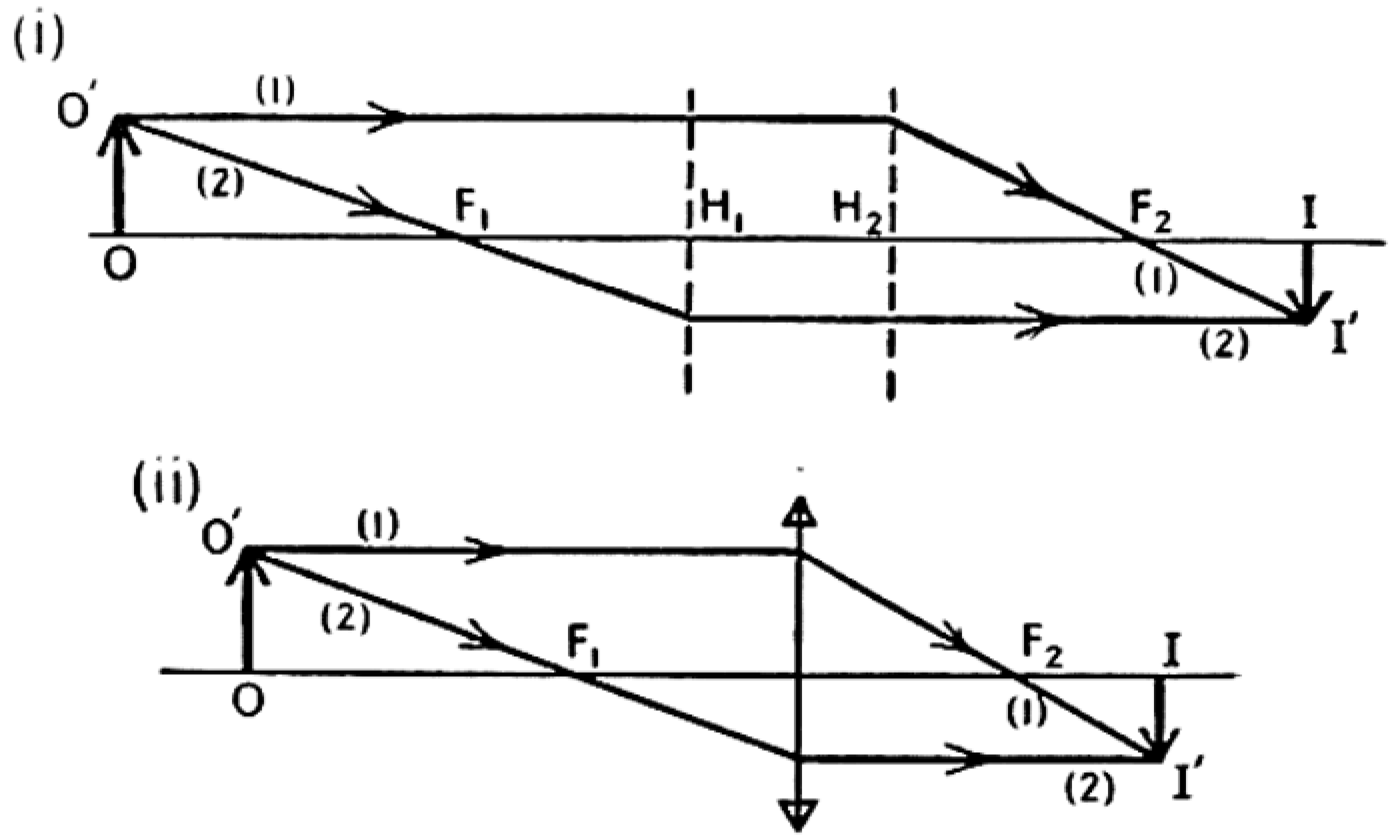


FIG. 616

When this is done and a comparison is made between Fig. 616 (i) and the corresponding ray diagram for a single thin converging lens (Fig. 616 (ii)), it is clear that we can regard the action of the combination on the incident light as being identical with that of a single thin lens, except that the plane at which the deviation is supposed to occur in the case of the thin lens becomes in the combination a space bounded by the two principal planes of the combination. In other words, the similarity between the rays outside the principal planes of the combination and the corresponding ones on either side of the single thin lens suggests that we could apply the thin-lens equation to the combination and write

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

where, as already explained, f is the position of either of the principal foci of the combination with respect to its appropriate principal plane, u is the position of the object relative to the first (or object) principal plane and v is the position of the image relative to the second (or image) principal plane. The fact that the simple lens equation can actually be used in this way can be verified by a geometrical proof from a ray diagram, or by

an algebraic treatment beginning with the application of the thin-lens equation to each component of the combination.

It follows also that we can use the magnification formula $-m = \frac{v}{u}$, Newton's formula (page 800), etc., thereby treating the combination as a single unit.

Cylindrical Lenses.—Suppose that a lens is bounded by two faces which are portions of cylindrical instead of (as we have hitherto supposed) spherical surfaces, and that the axes of the cylinders are parallel to each other. Fig. 617 (i) shows an example of such a lens in which the cylindrical

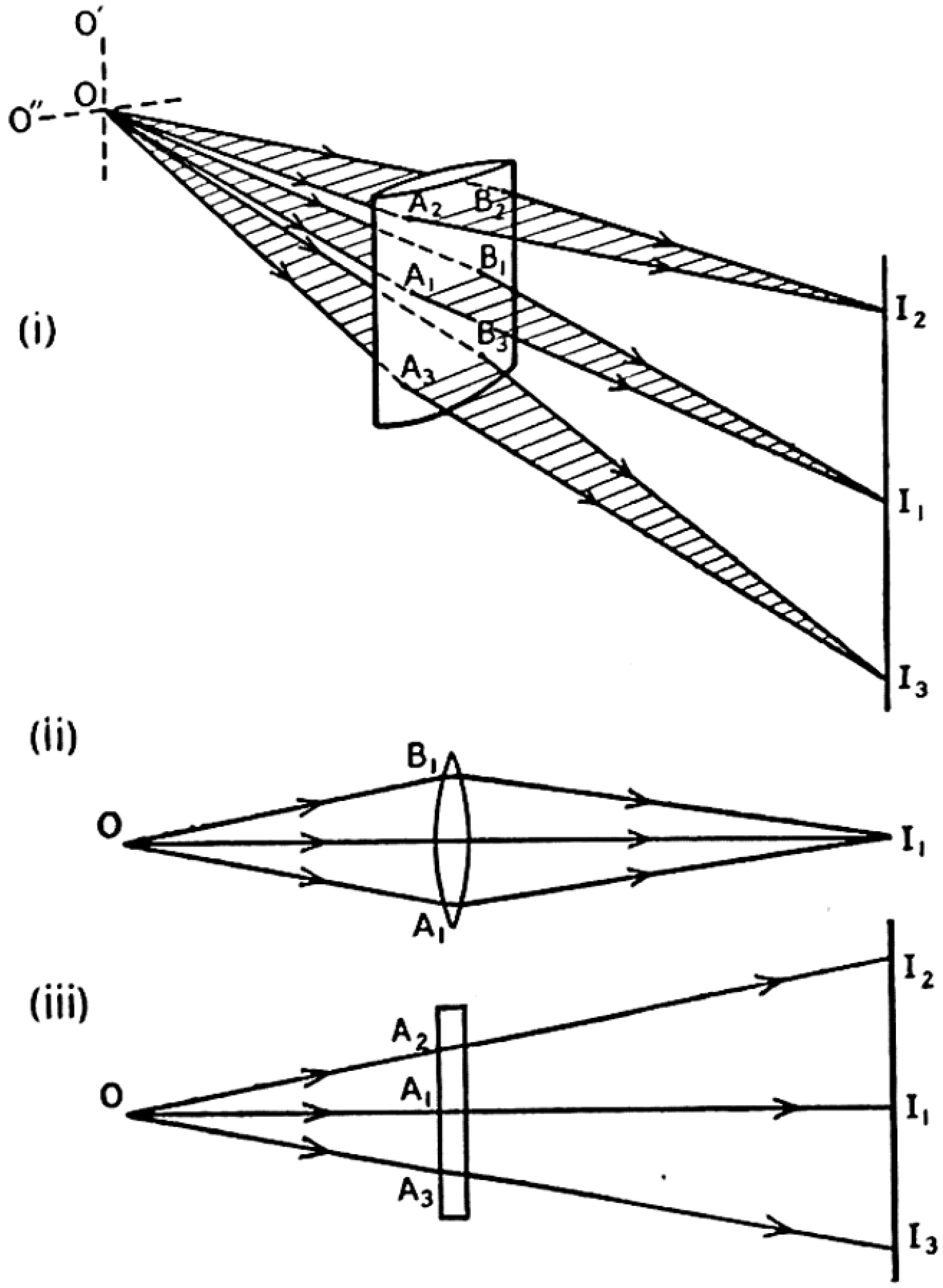


FIG. 617

surfaces are convex to the air. Consider the action of the lens on light from a point object situated at O. Incident rays such as OA_1 and OB_1 lying in the plane perpendicular to the axes of the surfaces will be refracted in exactly the same way as if the surfaces were spherical and will give rise to an image I_1. Seen from above, the ray diagram (Fig. 617 (ii)) is

similar to many that we have had in connection with spherical lenses. Rays like OA_2 and OB_2 which lie in a plane inclined to OA_1B_1 are also refracted in a similar way but are not made to converge to I_1 as they would be with a spherical lens, because, looked at from the side (Fig. 617 (iii)), the faces of the lens are parallel to each other and are not curved in this direction, so that no deviation occurs in the plane containing the axes of the surfaces. The rays continue in the plane OA_2B_2 after passing through the lens and are brought to a focus at I_2. Thus the image of the *point* O is a *line* I_2I_3 parallel to the axes of the lens surfaces. If the object is a line or illuminated slit like OO′ parallel to the axes its image is also a line parallel to the axes, each point on the object producing a line like I_2I_3 in the same place. But if the object is a line or slit like OO″ at right angles to the axes, each point on it will produce a line image like I_2I_3, and these images will be side by side and not superimposed, so that no true image is formed—only a patch of light. Evidently if the object is a crosswire or wire gauze with one set of wires parallel to the axes of the surfaces and the other set at right angles, the image will consist only of the former set. If two cylindrical lenses of equal power are placed in contact with each other, the axes of their cylinders being at right angles to each other, *i.e.* the lenses are "crossed," then the performance of the combination is exactly the same as that of a spherical lens of the same power as either of the cylindrical ones. Or again, a lens will behave as a spherical lens if it has two equal cylindrical surfaces with mutually perpendicular axes.

EXAMPLES XLV

1. Obtain from first principles a formula from which the focal length of a thin lens can be calculated in terms of the refractive index of the material and the radii of curvature of the two faces.

A double-convex lens, the radius of curvature of each face being 20·0 cm., is made of glass of which the refractive indices for red light and violet light are 1·643 and 1·685 respectively. A small incandescent lamp is placed on the axis of the lens at a distance of 35 cm. Show that the separation of the images formed in the red and the violet constituents of the light is approximately 3 cm. (O.H.S.)

2. Why is a sign convention employed in geometrical optics? State the convention you employ.

It is desired to form an image of a small object on a screen 45 in. from the object. This is done by placing a converging lens *A*, of focal length 4 in., at a distance of 6 in. from the object, and a second lens *B* between *A* and the screen 9 in. from *A*. Find the nature and focal length of *B* and the magnification of the image. (L.I.)

3. Explain the change in the focal power of a glass lens when immersed in water instead of air. If two watch-glasses of 20 cm. radii of curvature are cemented together to form an air-lens for use under water, what is its focal power when so used? Draw a diagram to show the behaviour of an incident parallel beam of light after its passage through the lens. (Assume that the refractive index of air-glass is $\frac{3}{2}$ and of air-water $\frac{4}{3}$.) (L.Med.)

4. How does the power of a lens depend on the curvature of its surfaces and on the refractive indices of the materials of which it is made and in which it is immersed?

An object is placed 25 cm. from a thin bi-convex lens. The radii of curvature of its surfaces are 20 and 25 cm. and the refractive index of its material is 1·5. Find the position, type and linear magnification of the image formed when the object and lens are (i) in air, (ii) immersed in a liquid of refractive index 1·66. (L.Med.)

5. A screen S_0, having an aperture across which a wire is mounted, is placed at an axial distance of 30 cm. from a thin converging lens C, and a real image of the wire with linear magnification 3/2 is focused on another screen S_1. If a convex spherical mirror M is now placed between S_1 and C and 21 cm. from the latter, an image of the wire is focused on S_0. What is the radius of curvature of M?

S_0, S_1 and C remaining unmoved, M is now removed and a thin diverging lens D and a plane mirror P are suitably placed between C and S_1. An image of the wire is focused on S_0 when the distance CD is 18 cm. What is the focal length of D?

S_0 and P remaining unmoved, C and D are now mounted coaxially in contact. To what position relatively to S_0 must the combination CD be moved so that the image of the wire is again focused on S_0?

In each of your three answers include a carefully drawn ray diagram. (L.Med.)

6. An object is placed 8 in. from a convex lens and the image formed is 24 in. from the lens. A plane mirror is then placed 6 in. from the lens on the side remote from the object. Determine the position of the image formed by the light reflected back through the lens if the first image is (*a*) real, (*b*) virtual. (L.I.)

7. Show that, when all the angles concerned are small, the deviation produced by a prism of refracting angle A is $(n-1)A$. Hence deduce a relation connecting the distances of an object and its image from a thin lens with the radii of curvature of the surfaces of the lens and the refractive index of the glass.

The radii of curvature of the surfaces of a thin convergent meniscus lens are equal two-ninths and three-ninths of its focal length respectively. What is the refractive index of the glass? (L.I.)

8. Derive an expression connecting the distances of an object and its image from a concave spherical mirror of small aperture with the radius of curvature of its surface.

An object is set up 8 cm. in front of a convex mirror of radius of curvature 20 cm. Where must a strip of plane mirror be placed so that the images of the object in the mirrors are in the same plane?

An object is placed $16\frac{2}{3}$ cm. from a convex lens of focal length 10 cm. and a real image is produced. The convex mirror ($r = 20$ cm.) is placed 20 cm. from the lens to intercept the light transmitted to it through the lens. Where will the image, formed by reflection at the mirror, be formed?

Where should the mirror be placed so that the *final* image of the source is in the same plane as the source? (L.A.)

9. Give an account of two different purposes for which concave lenses are employed.

A luminous object, $\frac{1}{4}$ in. long, is placed at right angles to the axis of a concave spherical mirror of 4 in. focal length and $4\frac{2}{3}$ in. from its surface. After reflection from the mirror the light passes through a concave lens of 10 in. focal length placed 23 in. from the mirror and coaxial with it and the object. Determine the length of the real image obtained. (L.I.)

10. Distinguish between real and virtual images.

An object is placed on the axis of a converging lens at a distance x from that principal focus which is situated on the same side of the lens as the object. Derive an expression for the transverse linear magnification of the image in terms of x and the focal length of the lens, and show that it has the same numerical value whether the image be real or virtual.

At what distances from a convex lens of focal length 15 cm. will an object give rise to an image having a transverse linear magnification of 3? (L.Med.)

11. What is meant by the *focal length* of a lens?

A real image of the sun is produced by a convex lens, the distance of the image from the lens being 20 cm. If a concave lens is placed in contact with the convex one, the image moves through 16 cm. and is still real. Calculate (*a*) the focal lengths of each of the lenses, and (*b*) the distance of the final image from the concave lens if the lenses are separated by 11 cm., the convex one being nearer to the sun. Explain the sign convention which you use. (L.I.)

12. A lens of small aperture forms an image of a luminous object on a screen placed 20 cm. from the lens. A glass cube of 5 cm. edge is then placed between the lens and the screen so that two of its faces are at right angles to the axis of the lens. If the refractive index of the glass is $\frac{5}{3}$, how must the screen be moved so that the image may again be in focus? Prove any formula used. (L.I.)

13. Deduce an expression for the focal length of a lens in terms of u and v, the object and image distances from the lens.

A lens is set up and produces an image of a luminous point source on a screen 25 cm. away. If the aperture of the lens is small, where must the screen be placed to receive the image when a parallel slab of glass 6 cm. thick is placed at right angles to the axis of the lens and between the lens and the screen, if the refractive index of the glass is 1·6? Deduce any formula you use. (L.I.)

14. Parallel light incident on a glass lens of focal length f is reflected twice internally before emerging on the far side of the lens. At what distance from the lens is the emergent light brought to a focus? (Refractive index of glass $=\frac{3}{2}$.)

Chapter XLVI

DEFECTS OF LENSES*

1. SPHERICAL ABERRATION AND KINDRED DEFECTS

Spherical Aberration.—The formulæ we have proved and used in connection with spherical lenses up till now have, as has been frequently emphasized, been confined to cases in which all rays are paraxial or nearly so. This is because we have supposed that the angles of incidence and refraction have been sufficiently small for us to replace the sines in the exact form of Snell's law, namely

$$n_1 \sin i_1 = n_2 \sin i_2$$

by the circular measure of the angles, thus writing

$$n_1 i_1 = n_2 i_2$$

We shall now describe the effect which this approximation has on the simple case of the refraction of a parallel axial beam by a converging spherical lens (Fig. 618).

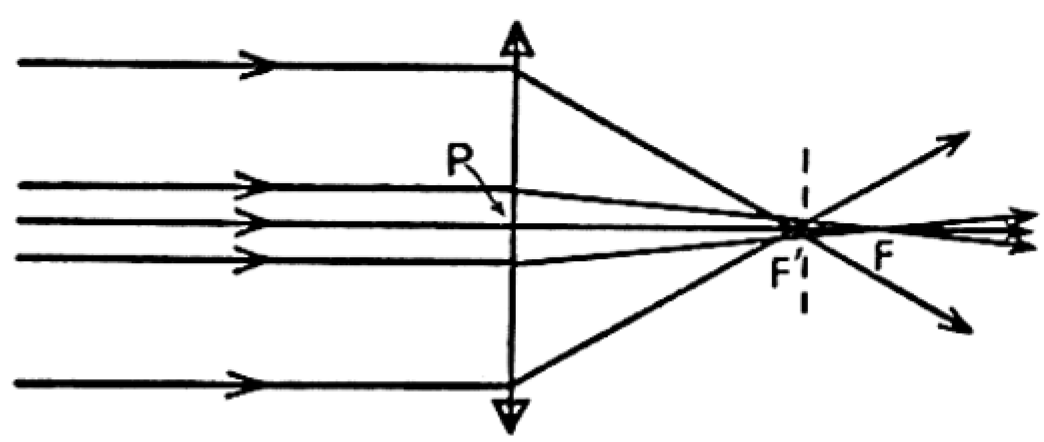


FIG. 618

Suppose that rays of light parallel to the principal axis of a converging lens arrive from a very distant point object. If the incident rays are confined to those near the axis by placing a stop, *i.e.* a screen with a small hole in it, in contact with the lens, the centre of the hole being on the axis of the lens, then an image will be formed at F, the position of which can be calculated from the formula for the focal length of a lens which was derived on page 789 for paraxial conditions. If, however, incident rays are confined to those striking the lens at a considerable distance from the axis by using a screen with an annular zone cut out of it, the refracted rays are deviated so as to come to a focus at some point such as F′ which is nearer to the lens than F. The distance FF′, which is called the **axial spherical aberration** of the particular zone of the lens through which the outer rays are passing, increases (*i.e.* F′ moves nearer to the lens) as the radius of the zone is increased, reaching a maximum when the rays just pass through the edge of the lens—the so-called marginal rays.

* Certain parts of this chapter are somewhat more advanced than the general standard of this book. They are marked with an asterisk (*).

It is obvious, therefore, that a clear image of a small object such as an illuminated pinhole cannot be formed if all the possible incident rays are allowed to fall on the lens at once, since light from any one point on the object is nowhere brought to a single focus. The position in which a screen can then be placed so as to receive an image which is most nearly what it should be is shown by the vertical dotted line in Fig. 618. This smallest patch of light is called the **circle of least confusion.**

Spherical aberration could of course be minimized by using a stop to confine the refraction to the small central zone of the lens although, in the simple case depicted in Fig. 618, it is clear that this would reduce the illumination of the image, which is a serious disadvantage in, say, a camera lens. It is also possible to eliminate spherical aberration by modifying the surfaces from spherical to an appropriate shape, which can be calculated, but the elimination is complete only for one object distance.

As we have seen, the form of Snell's law in which the sines of the angles of incidence and refraction are replaced by the angles themselves is only a first approximation. The full expression for the sine of an angle θ in terms of θ in radians is really the sum of an infinite series of terms of which θ is the first. Thus

$$\sin\theta = \theta - \frac{\theta^3}{6} + \frac{\theta^5}{120}, \text{ etc.}$$

successive terms decreasing rapidly when θ is a small fraction of a radian. If, as a second approximation, we substitute $\left(\theta - \frac{\theta^3}{6}\right)$ for $\sin\theta$ in Snell's law we have

$$n_1\left(i_1 - \frac{i_1^{\,3}}{6}\right) = n_2\left(i_2 - \frac{i_2^{\,3}}{6}\right) \quad . \quad . \quad . \quad . \quad (1)$$

This equation will represent the refraction of the marginal rays more accurately than the approximate equation

$$n_1 i_1 = n_2 i_2 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

which applies accurately enough to the central rays and gives the position F for the image in Fig. 618. As the incident rays get further from the axis, the curvature of the lens surfaces causes i_1 and i_2 to increase beyond the limits within which the approximate equation is applicable, and the refraction, which brings the marginal rays to a focus at F′, is more truly represented by equation (1), containing the first correction, than by equation (2). We can therefore see in a semi-quantitative way that the order of magnitude of the aberration as represented by F′F is related to the order of magnitude of the quantities $\frac{i_1^{\,3}}{6}$ and $\frac{i_2^{\,3}}{6}$. When the angle of incidence (and therefore the angle of refraction) of a ray is increased, either by increasing the

curvature of the surface upon which it is incident or by moving the ray further from the axis, the angles i_1 and i_2 are increased, and the aberration which is related to the cube of these angles is increased by a greater factor. For instance, if the angle of incidence is multiplied by a factor of 2 the term $\frac{i_1^3}{6}$ is multiplied by 8.

The deviation at any one face of a lens increases with the angle of incidence. Therefore we can say that the aberration, which increases rapidly with angle of incidence, also changes in a similar way with the deviation. Consequently, in order to minimize the aberration associated with a given total deviation produced by a given zone of a given lens it is necessary to share the deviation equally between the two faces, because if, starting with equal deviations, the deviation at, say, the first face is increased and that at the second decreased by the same amount, then the increase of aberration at the first face will exceed the decrease of aberration at the second. As a simple application of this principle we may consider the case of a parallel beam falling on a plano-convex lens. In Fig. 619 (i)

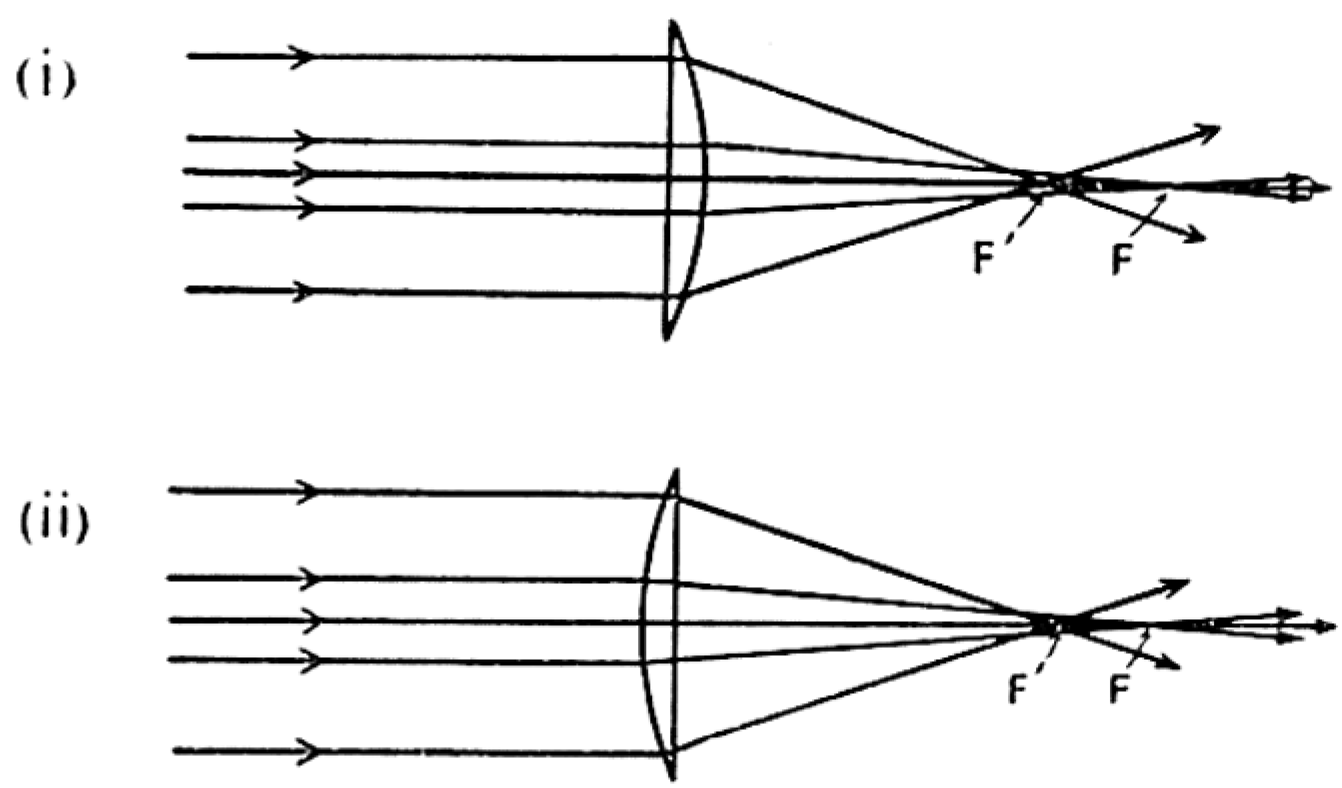


Fig. 619

the light is incident normally on the plane surface through which it passes without deviation. The whole of the deviation due to the lens occurs at the curved surface and consequently there is considerable spherical aberration. If, however, the lens is turned round as in Fig. 619 (ii), the deviation is more equally shared between the surfaces and the aberration is less. Thus the amount of aberration produced by a lens depends, in general, on the direction in which the light passes through it. It is evident also that the aberration depends upon the position of the object. We have seen that the aberration is large in Fig. 619 (i) when the object is very distant, but as the object is brought nearer, more deviation occurs at the first face and less at the second, thus reducing the aberration. This explains why, as has already been mentioned, a lens can be corrected for spherical aberration for only one position of the object,

As a general rule it may be said that if we are given a plano-convex lens with which to form an image of an object situated at a *moderate* distance from the lens then the deviation will be more equally shared, and the aberration will be less, if the plane surface faces the object. If the object is very distant, however, the curved surface should face it.

Coma.*—There are other defects from which the image formed by a lens suffers, and they are concerned with the formation of an image of a point which is *not* situated on the axis of the lens. Fig. 620 illustrates the refraction of rays from a non-axial point O′ by a lens. Spherical

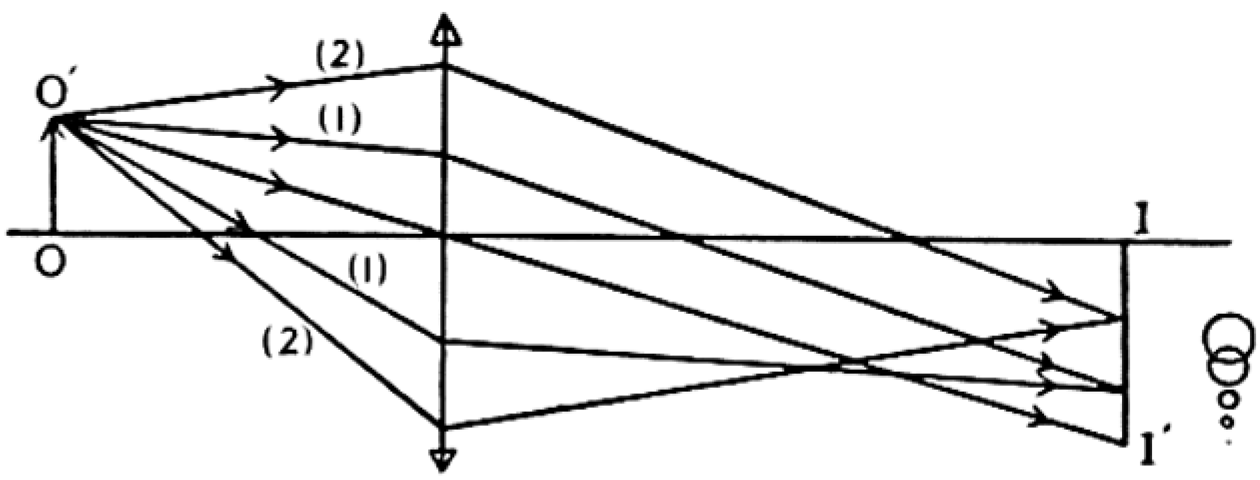


FIG. 620

aberration is supposed to be absent, that is to say the rays from O′ through the centre of the lens come to a focus in the same plane (II′) as do all the other pairs of rays such as (1), (1) and (2), (2). But unless further steps are taken, a single point image of O′ is not formed because, as can be seen from the diagram, rays passing through the different zones of the lens come to a focus at different points in the plane II′, the foci moving from I′ towards the axis as the radius of the effective zone is increased. This defect occurs because, owing to O′ being off the axis, the angle of incidence is different for the two rays of any one of the pairs such as (1), (1) and (2), (2), and this disparity and the consequent displacement of the image from I′ increases with the width of the zone.

Fig. 620 shows only the refraction of incident rays which lie in the plane containing O′ and the axis. When, for any one zone, the other (so-called "skew") rays are considered it is found that each zone produces a circle of light, the diameter of the circles increasing as we go from I′ towards the axis. The series of circles formed by the various zones is indicated on the right-hand side of the drawing. Thus when the non-axial point O′ is, say, an illuminated pinhole in a screen and a screen is placed at II′ to receive the image, the latter consists of a patch of light of the shape indicated by the superposition of the circles, the intensity being greatest at the pointed end I′ of the patch. The appearance resembles that of a comet with a tail, hence the name "coma."

It will be seen from Fig. 620 that coma can be reduced by using a

* See footnote, page 815.

small stop close to the lens to confine the rays forming the image to those passing through the centre of the lens. Such a stop, therefore, reduces the effects of both spherical aberration, which draws the image out along the axis of the refracted pencil, and coma, which extends the image sideways.

Astigmatism.*—This defect of lenses concerns the nature of the refracted beam when the object is situated some distance from the principal axis. We have seen that the elimination of spherical aberration and coma by the use of a central stop would cause a slightly non-axial point object to give rise to a single point image, but in fact the inherent asymmetry of even a narrow pencil of incident rays with respect to the

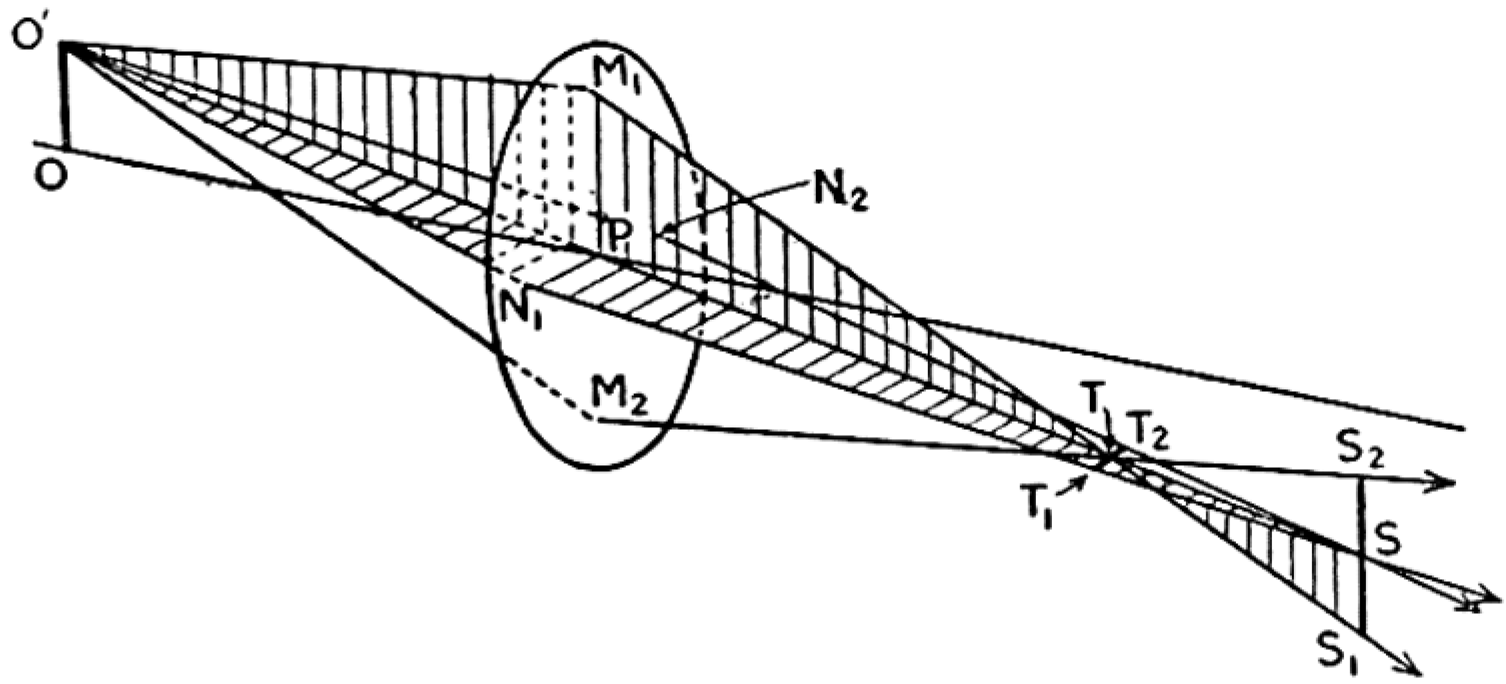


FIG. 621

lens prevents the refracted pencil from coming to a point focus when the object-point is at an appreciable distance from the axis. The corresponding effect in mirrors has already been described on page 736.

Fig. 621 shows the type of refracted beam which is produced by a non-axial point object. For the purpose of specifying the character of the refracted beam we define two planes. The **meridional plane** contains the object point O′ and the principal axis of the lens. The incident rays $O'M_1$ and $O'M_2$ lie in this plane. The plane in which O′ and P lie and which is perpendicular to the meridional plane is called the **sagittal plane,** and the rays $O'N_1$ and $O'N_2$ lie in this plane. It will be seen from the drawing of the refracted rays that the meridional rays are brought to a focus at T on O′P produced, and that *all* the rays, both meridional and sagittal and those in intermediate positions, pass through the straight line T_1T_2 which is at right angles to the meridional plane. Similarly all the sagittal rays are brought to a focus at S and all the rays pass through the straight line S_1S_2 perpendicular to the sagittal plane. Thus if a luminous point object such as a pinhole is placed at O′, and a screen is held at right angles to the auxiliary axis of the lens (O′P produced) and moved to and from the lens, two positions will be found for which a sharply focused line is formed on the screen. In moving the screen from

* **See footnote, page 815.**

one of these positions to the other the patch of light changes into a widening ellipse which becomes a circle in one position (the circle of least confusion) and then passes through a series of narrowing ellipses until it becomes linear again. These changes are illustrated in Fig. 622. The line T_1T_2 is called the **tangential focal line,** and S_1S_2 is the **sagittal focal line.** The reason for these names is explained later. It should be noticed that pairs of rays which do not lie in either the meridional or the sagittal plane do not come to a focus anywhere. Imagine that a pair of rays such as $O'M_1$ and $O'M_2$ lying initially in the meridional plane are rotated so as eventually

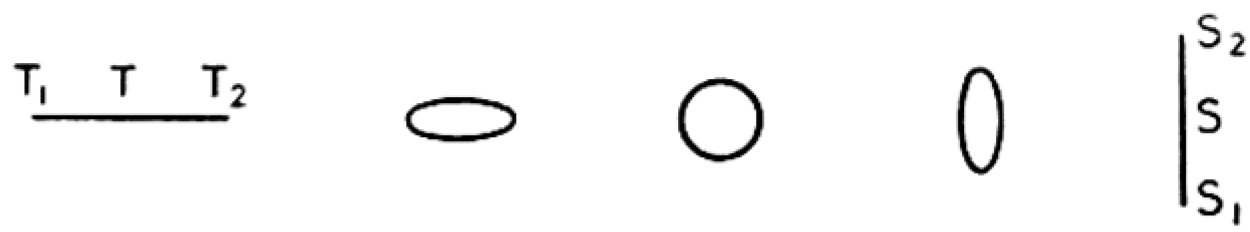


FIG. 622

to become sagittal rays $O'S_1$ and $O'S_2$. This causes the points at which the two refracted rays cross T_1T_2 to separate and to move out from T to opposite ends of the focal line, while the points at which the same rays cross S_1S_2 move in from the ends until they coincide at S. The effective focal length of the lens for both the tangential and the sagittal focal lines is less than its focal length for an object situated on its axis, and as the inclination of the auxiliary axis O′P to the principle axis OP increases, the focal lines get nearer to the lens, *i.e.* the effective focal length for each decreases.

We now consider the formation of images by other points near O′ in the line OO′ (Fig. 621), which is perpendicular to the principal axis OP. Evidently as the point object moves slightly from O′ towards O the focal line S_1S_2 is moved towards the axis along its own direction while T_1T_2 is moved towards the axis in a direction perpendicular to itself. Therefore the only clear image of a short line object lying along OO′ near to O′ will be a line in the neighbourhood of S_1S_2. If the line object extends for a considerable distance along OO′ the image will cease to be straight, because as a point object approaches the principal axis the focal lines move away from the lens.

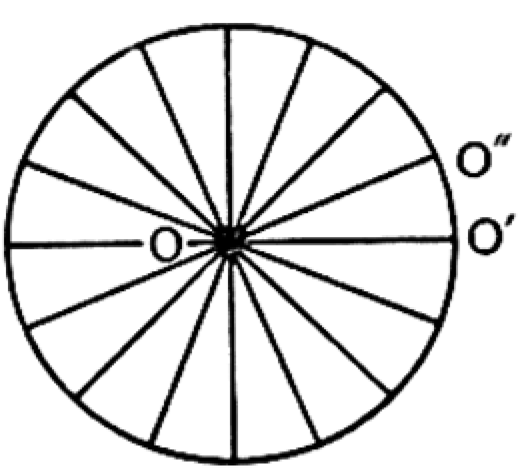


FIG. 623

It is interesting to consider the effect of using as an object a kind of spoked wheel (Fig. 623) with its centre at O on the principal axis. Referring to Fig. 621 once again, it is evident that the portions of the spokes near the points like O′, O″, etc. on the circumference will, according to what has been said in the last paragraph, each form a sagittal focal line like S_1S_2, and these lines will be inclined to each other in the same way as the spokes, for if OO′ in Fig. 621 is rotated through a given angle about the principal axis, the focal lines due to O′ rotate through the same angle. Thus when a screen is placed on the other side of the lens we can

pick up the images of the outer ends of the spokes at a certain distance from the lens, and as the screen is moved away from the lens from this position we see the images of the portions of the spokes near the centre and eventually of the centre itself. This is why the word "sagittal" is used. For a given position of the screen, images of short portions of the spokes are seen, each pointing like an arrow (Latin *sagitta*) towards the centre.

If we consider the effect on the tangential focal line due to O′ of rotating OO′ about the principal axis, we see that T will describe a circle (also around the axis) and that in any one position the focal line T_1T_2 is a tangent to this circle (hence the term "tangential"). Thus the complete rim of the wheel will give a circular image of itself in the plane of the tangential focal line. The spokes are not in focus at this plane and, conversely, the rim is not in focus when the screen is placed to receive an image of portions of the spokes.

Curvature of the Field.*—In the absence of spherical aberration, coma and astigmatism, all the rays which strike the lens from a non-axial point object, such as O′ in Fig. 624, are brought to the same focus I′.

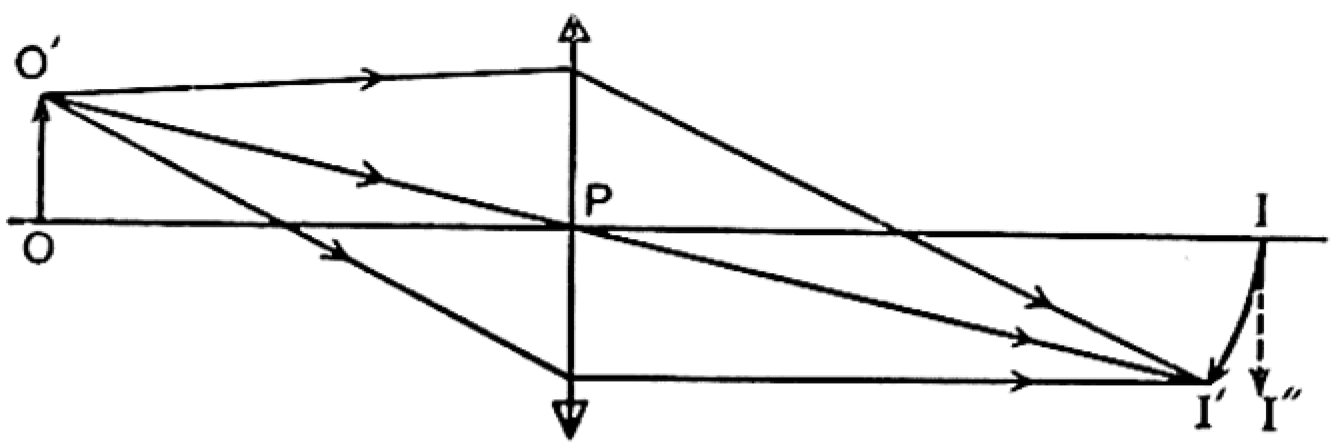


FIG. 624

In discussing the theory of thin lenses when paraxial conditions are satisfied we have said (page 795) that if I is the image of the axial point O (OP and IP being related to the focal length by the thin-lens equation), then an object OO′ perpendicular to the principal axis gives an image II″, also perpendicular to the axis. We can only expect this to be true, however, when the aperture of the lens and the distance of O′ from O are very small compared with distances like OP and IP. The effect of a departure from these conditions can readily be deduced. The line O′PI′ is an auxiliary axis of the lens and, neglecting astigmatism, we can apply the equation

$$\frac{1}{I'P}+\frac{1}{O'P}=\frac{1}{f}$$

to object and image points such as O′ and I′ on this axis, f being the focal length of the lens for objects and images on the principal axis. Therefore

$$\frac{1}{I'P}+\frac{1}{O'P}=\frac{1}{IP}+\frac{1}{OP}$$

* See footnote, page 815.

The distance O′P is evidently greater than OP, which means, according to the last equation, that I′P is less than IP. Therefore the line joining I to I′ (*i.e.* the image of OO′) is a curve, as shown in Fig. 624, and a screen would have to be of this shape in order to obtain a clear image. It would be impossible for every point on the image to be in focus at the same time if a flat screen were used. As has already been mentioned, this defect of curvature of the field, being due to the different distances of the various parts of a flat object from the centre of the lens, would occur even when the other defects are absent. When astigmatism is present and is superimposed on curvature, each point on the object which would have given a single point on the curved image now gives a tangential and a sagittal focal line each of which is nearer the lens than the point image. Thus the astigmatism not only enhances curvature but causes a blurring of the image, so that it is not possible to form a clear image even on a curved screen. The best position and shape of the screen is, as shown in Fig. 625, a compromise, and is known as the **surface of least confusion**

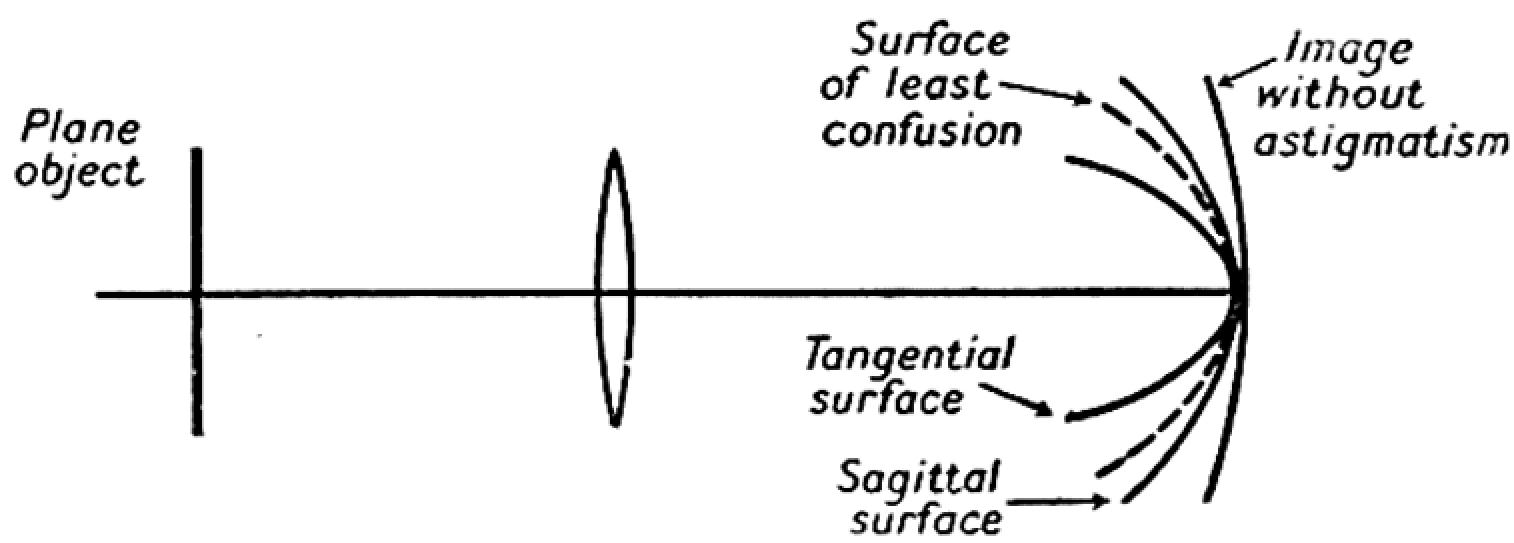


FIG. 625

which is situated somewhere between the surfaces on which the tangential and sagittal images are formed.

Distortion.*—When all the defects so far mentioned are absent, a plane object produces a plane image, every point on which is a clear image of the corresponding point on the object. But there is still the possibility that the image is imperfect because the magnification may not be the same all over the image—the ratio of the distance between a pair of points on the image to that between the corresponding points on the object may vary with the positions of the points. This defect is known as distortion. Its character is influenced by the way in which the rays are restricted to different parts of the lens by a stop.

In Fig. 626 a stop is interposed between a plane object OO′ and a converging lens. The centre of the stop is at S on the principal axis, and we consider rays passing through this point from points on the object such as O′ and O″. In the absence of other defects, the images of these points lie in the plane II′ at I′ and I″ respectively. One of the rays by which each of these images is formed is shown in Fig. 626 passing

* See footnote, page 815.

through the centre of the stop S. We now compare the paths of these two rays from S onwards. We can evidently regard the centre of the stop S as a point object, and the drawing shows two rays (the original ones from O′ and O″) from S passing through the lens. It has been mentioned before that spherical aberration can be eliminated only for one

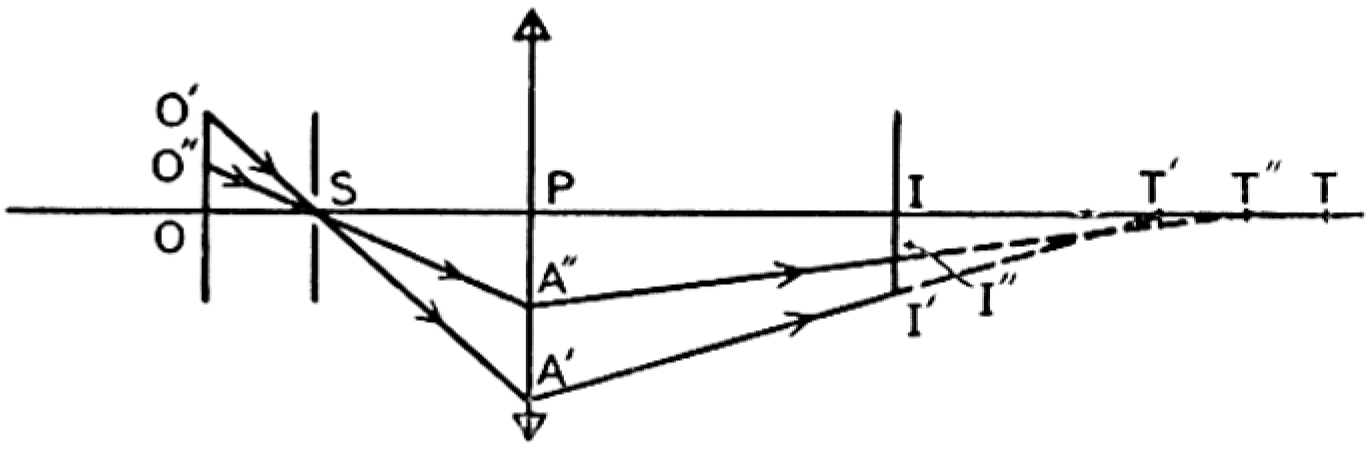


FIG. 626

position of the object. Suppose that it is absent for the object OO′ and its image. It will however be present for the rays forming the image of S, so that, according to the characteristic feature of this aberration, the ray passing through the outer zone of the lens at A′ will cross the axis (at T′) nearer the lens than the other ray through A″, which crosses it at T″.

The defect we are discussing is effectively spherical aberration of the image of S. If this aberration had been absent (say by using only the central zone of the lens), the image of S would have been at some point T beyond T′ and T″, that is to say I′ would have been on the straight line A′T and I″ on A″T, as shown in Fig. 627.

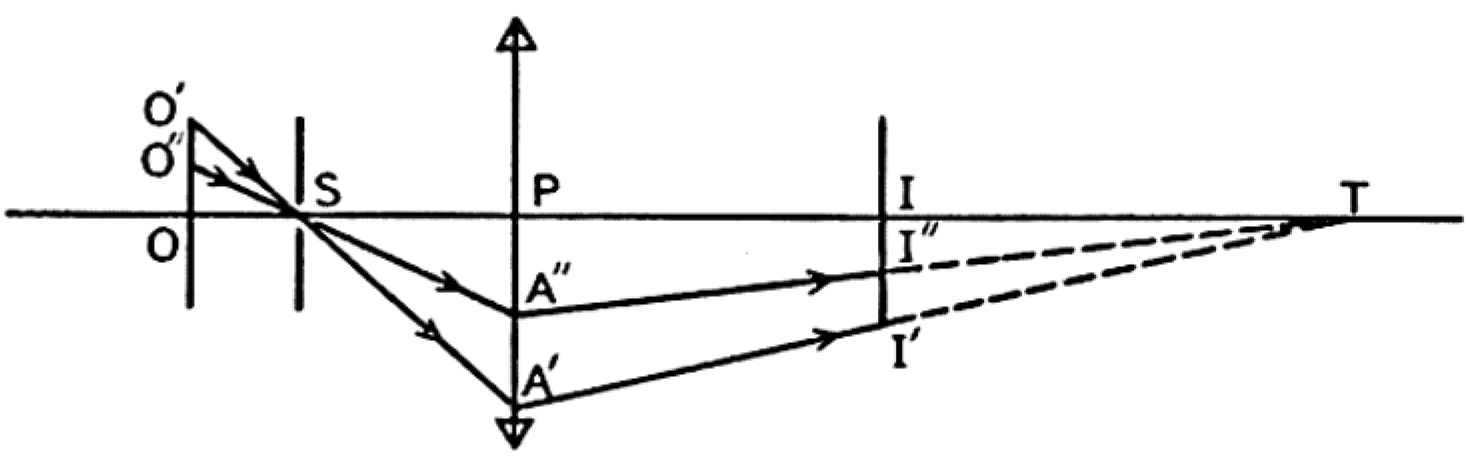


FIG. 627

It can be seen from the drawing (Fig. 626) that, on account of the positions of T′ and T″ due to spherical aberration, the actual image lengths II′ and II″ are both shorter than they would be if T′ and T″ both coincided at T. Therefore the magnification in each case is less than it would be if the image were perfect. Furthermore, the fact that T′ is nearer the lens than T″ makes the reduction of the magnification greater for II′ than for II″. Therefore the image is distorted because the magnification of the distance between the axial point and a point some distance from the axis diminishes as the non-axial point gets farther from the axis.

The effect of this distortion on the shape and size of the image can be described by reference to a square plane object with its centre on the

principal axis. In the absence of this type of distortion its image would be a square such as that shown dotted in Fig. 628. The general reduction of magnification mentioned above makes the image smaller than this, but it does not retain its square shape. The distance from the centre of the image I to a corner is greater than from I to the midpoint of a side. The former distance is therefore less magnified than the latter, and the result

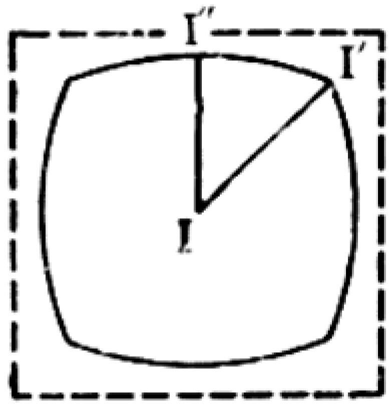


FIG. 628

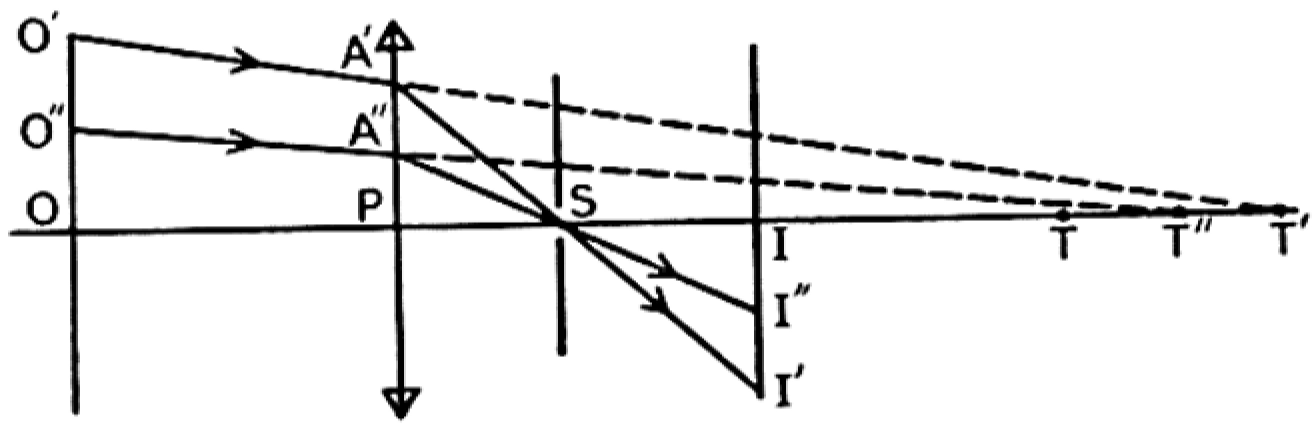


FIG. 629

is a shape like that shown in Fig. 628. For obvious reasons this is called **barrel distortion.**

When the stop is placed after the lens, between the lens and the second principal focus, the ray diagram corresponding to Fig. 626 is Fig. 629. To study the effect of spherical aberration we again consider the centre of the stop S as an object. In this case we must suppose the path of the light to be reversed. If there were no spherical aberration the (virtual) image of S would be at T, say, but aberration causes the rays SA″ and SA′ to be deviated more than they would if it were absent, and the effect on SA′ is greater than on SA″; hence the relative positions of T, T′ and T″. As in the previous case, we compare the sizes of II′ and II″ with what they would have been if T had been the one and only image of S. If this had been the case, the incident rays O′A′ and O″A″ would have been directed towards T, so that A′ and A″ would have been nearer P than they are in the figure and consequently I′ and I″ would have been nearer I. The magnification is therefore greater than when there is no distortion, and it can be seen that the magnification of the greater length II′ is larger than for II″. A square object is therefore distorted into the shape shown in Fig. 630. This is known as **pincushion distortion.** The student can consider for himself the case of a stop beyond the second principal focus. The argument then involves a real image of S on the object side of the lens. This case also results in pincushion distortion.

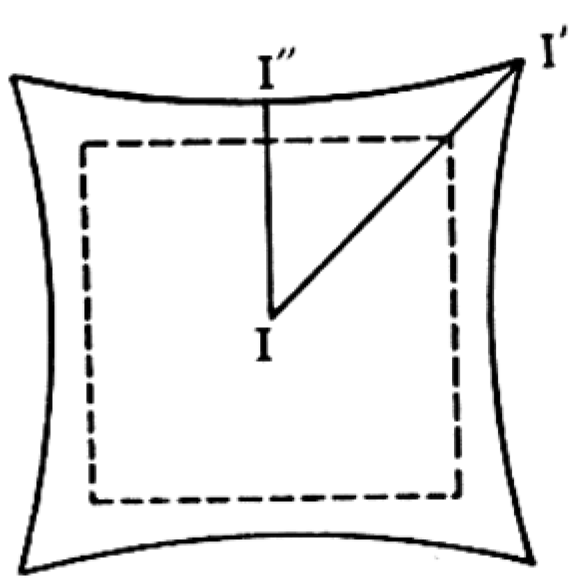


FIG. 630

The effect of the position of the stop on the type of distortion produced can be clearly demonstrated practically when an illuminated wire gauze is used as the object.

The fact that the distortion changes from one type to the other when the stop is transferred from one side of the lens to the other explains why distortion can be minimized by using two separated lenses with a single stop placed between them. The pincushion distortion produced by the first lens and the stop is counterbalanced by the barrel distortion due to the stop and the second lens. This arrangement also reduces coma.

2. CHROMATIC ABERRATION

The Effect of Dispersion on the Image.—The various defects of the image which have been discussed in Section 1 of this chapter are sometimes called monochromatic aberrations because they are present even when monochromatic light is used. We now come to **chromatic aberration,** which is due to the dispersion of light by the material of the lens and is therefore completely eliminated when monochromatic light is used.

The effect of dispersion when light of different wave-lengths is refracted by a thin prism is described on page 767. The deviation produced increases with decreasing wave-length, *i.e.* as we go from red to violet through the visible spectrum. A lens acts as a thin prism whose refracting angle varies from zero at its centre to a maximum at its margin. It is therefore clear that if a ray of white light is incident parallel to the principal axis on a converging lens, the light is analysed by the lens into its various colours and each coloured ray will cross the axis at a different point. The extreme violet and red refracted rays are shown in Fig. 631. The effect of dispersion is therefore to cause the focal length of a lens to vary with the wave-length of the light, the shortest focal length being for violet light and longest for red. Thus when a small object is emitting white light there is no one plane at which its image is perfectly distinct.

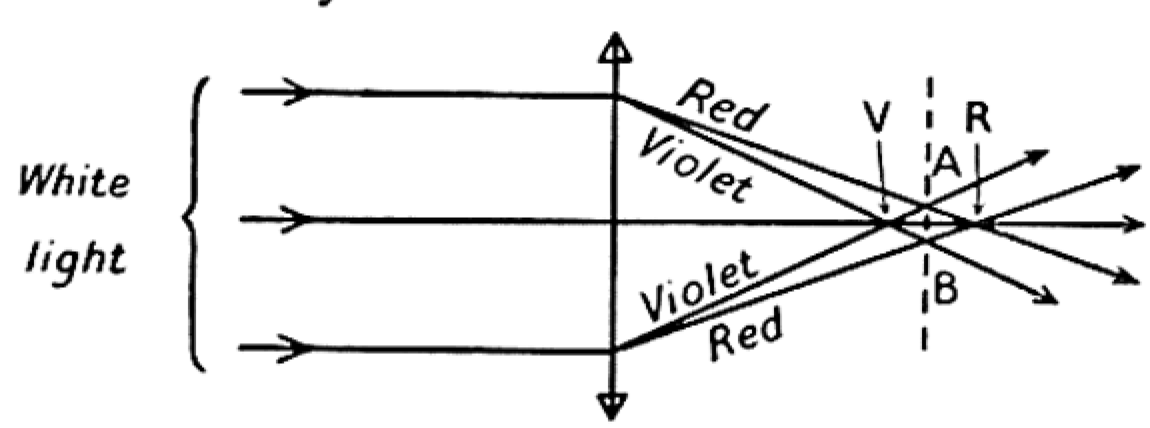


FIG. 631

It is clear from Fig. 631 that if a screen is placed in the position of the focus for violet light the edges of the image will appear coloured, the red being on the outside, while with the screen at the focus for red the edges will be violet. The red and violet will be superimposed at the edge of the image when the screen is in the dotted position, but this does not mean that the image is free from any colour effects because different pairs of colours are superimposed at other places on the screen. In any case the image will be blurred because it should be a point (if we suppose that the incident rays are all coming from the same point on the object),

and the smallest patch of light which can actually be obtained is the **circle of least confusion** AB. It is evident that if, in Fig. 631, the incident rays are brought nearer to the axis and the positions of the foci for the various colours are not thereby moved (*i.e.* if there is no spherical aberration), then the diameter of the circle of least confusion is proportionately reduced. Thus the blurring effect of chromatic aberration can be reduced by stopping down the lens.

A good demonstration of chromatic aberration can be made by magnifying its effect. This is done by the arrangement shown in Fig. 632.

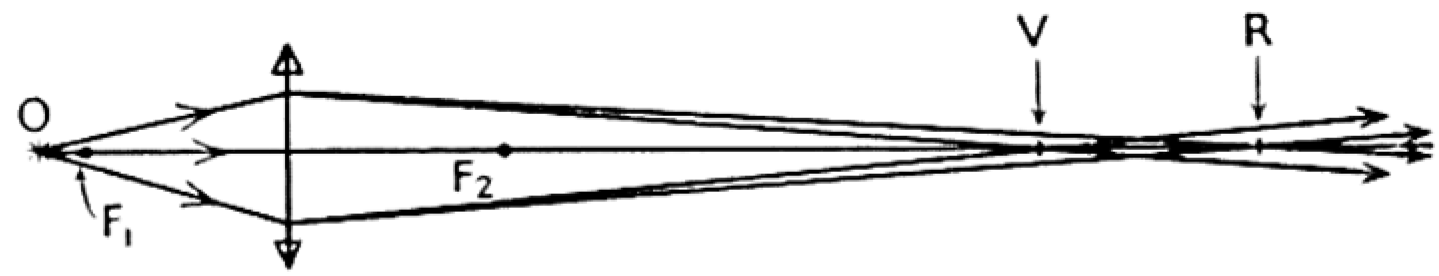


FIG. 632

An ordinary converging lens is placed so that O, a small bright white-light source, is only a short distance further from the lens than the first principal focus F_1. This is the condition for the formation of a highly magnified real image, so that there is a correspondingly large region in which the effects of chromatic aberration can be observed by placing a screen in various positions.

The effect of chromatic aberration when parallel light is incident on a diverging lens is shown in Fig. 633. The focal length is again shorter for the short wavelengths.

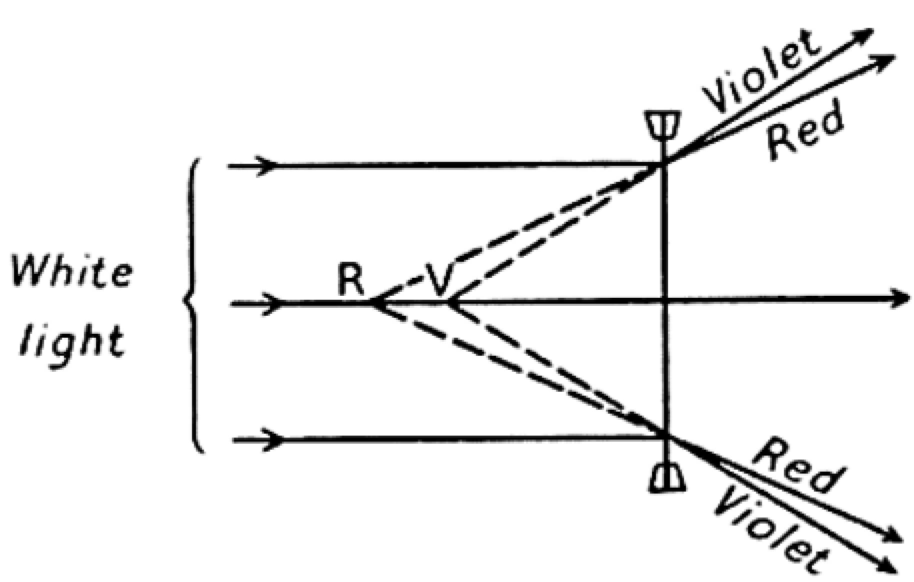


FIG. 633

In order to estimate the magnitude of chromatic aberration we shall consider an equi-convex lens of glass for which the refractive index is 1·50 for light of an intermediate colour such as yellow, while its values for red and violet are 1·49 and 1·51 respectively. Suppose that each face of the lens has a radius of curvature of 20 cm. Its focal length f_Y for yellow light is then given by

$$\frac{1}{f_Y}=(1{\cdot}50-1)\left(\frac{1}{20}+\frac{1}{20}\right)$$

so that

$$f_Y=\frac{10}{0{\cdot}5}$$

$$=20\ \text{cm}.$$

It is obvious that for red light we have

$$f_{\mathrm{R}}=\frac{10}{0{\cdot}49}$$

$$=20{\cdot}4 \text{ cm.}$$

and for violet light

$$f_{\mathrm{V}}=\frac{10}{0{\cdot}51}$$

$$=19{\cdot}6 \text{ cm.}$$

Thus the difference between the focal lengths for the extreme ends of the visible spectrum is 0·8 cm. The effect of this aberration can be considerably greater than that of spherical aberration and, although the falling off of the sensitivity of the eye at the ends of the visible spectrum makes the effect less troublesome than might at first appear, yet it is large enough to make it impossible to produce satisfactory refracting (*i.e.* lens) telescopes, microscopes, etc. unless steps are taken to reduce chromatic aberration.

The Correction of Chromatic Aberration.—On page 769, where the dispersion due to a thin prism is discussed, it is shown that dispersion is eliminated by placing two thin prisms of different materials in opposition provided that the ratio of the refracting angles satisfies equation (6) on page 769. The elimination is perfect for only two wave-lengths, however. The identical principle is applied to the elimination of chromatic aberration in lenses, and this is not surprising because a lens may be regarded as a thin prism whose angle varies from the centre to the margin.

We have seen (Fig. 631) that, with a converging lens, the violet rays of incident parallel white light are deviated towards the axis more than the red rays, and that with a diverging lens (Fig. 633) the violet is more deviated away from the axis. Therefore it should be possible to eliminate the chromatic aberration of a converging lens by following it with a diverging lens which produces an equal and opposite dispersion. As with prisms, so in this case, it is necessary that the lenses should be made of different materials, otherwise, by matching their dispersions, we shall also match their focal powers and the combination or "doublet" would have no resultant power.

Let the two colours for which the doublet is to be achromatized be the F (blue) and C (red) lines of the spectrum, and let the light of intermediate wave-length to which dispersive power is referred be the D (yellow) line. For the first lens let the refractive indices for these colours be n_{F}, n_{C}, n_{D} respectively. Then the dispersive power of its material (ω) is given by

$$\omega=\frac{n_{\mathrm{F}}-n_{\mathrm{C}}}{n_{\mathrm{D}}-1}$$

so that

$$n_{\mathrm{F}}-n_{\mathrm{C}}=(n_{\mathrm{D}}-1)\omega \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The power of the lens for the blue light (K_F) is given by

$$K_F = (n_F - 1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \quad . \quad . \quad . \quad . \quad (4)$$

which we shall write as

$$K_F = (n_F - 1)B$$

where B stands for the sum of the curvatures of the lens surfaces. For the C red light the power of the first lens is given by

$$K_C = (n_C - 1)B$$

For the second lens we can write

$$K_F' = (n_F' - 1)B'$$

and

$$K_C' = (n_C' - 1)B'$$

The condition that the combination shall be achromatic for the two extreme colours chosen is that it shall have the same focal power for the two colours, and since the focal power of a combination is the sum of the powers of the separate lenses, this condition can be written

$$K_F + K_F' = K_C + K_C'$$

or

$$K_F - K_C = K_C' - K_F'$$

so that, substituting for the powers, we obtain

$$(n_F - 1)B - (n_C - 1)B = (n_C' - 1)B' - (n_F' - 1)B'$$

$$\therefore \ (n_F - n_C)B = -(n_F' - n_C')B'$$

Hence by equation (3)

$$\omega(n_D - 1)B = -\omega'(n_D' - 1)B' \quad . \quad . \quad . \quad (5)$$

Now $(n_D - 1)B$ is the power of the first lens for the yellow D line and $(n_D' - 1)B'$ is the power of the second lens for the same colour. If we call these K_D and K_D' respectively, the condition for achromatism (equation (5)) becomes simply

$$\omega K_D = -\omega' K_D' \quad . \quad . \quad . \quad . \quad (6)$$

or, since power is the reciprocal of focal length,

$$\frac{\omega}{f_D} = -\frac{\omega'}{f_D'}$$

The negative sign indicates that the lenses have opposite powers, one being converging and the other diverging. Equation (6) shows that the lens with the numerically greater focal power, *i.e.* the one which is of the same type as the combination, must have the smaller dispersive power. Thus for a converging achromatic doublet the converging component has the smaller dispersive power.

As mentioned on page 770, crown and flint glasses are frequently used in combination for correcting for chromatic aberration. Typical values of n and ω for these glasses are given on page 770. The crown glass has the smaller dispersive power.

As with prisms, so with lenses, a combination of *two* different glasses gives achromatism in respect of *two* colours only. Three lenses would be required for the superposition of three colours, and so on.

In making achromatic doublets the two adjacent surfaces of the lenses are given the same curvature so as to fit together, and they are often stuck together by Canada balsam, which has a refractive index (1·55) near those of the glasses and so causes less loss of light by reflection at the lens surfaces than if there were an air film between them. The curvatures of the two outer surfaces of the doublet can be made so as to give the minimum spherical aberration in the conditions under which the lens is used. As stated on page 817, this aberration is minimized by sharing the deviation equally between the surfaces. Fig. 634 is a diagram of a typical achromatic doublet of crown and flint glass.

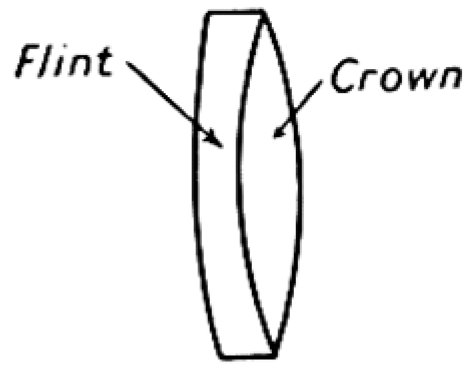


FIG. 634

Until it was known that various types of glass had different dispersive powers and so could be used to diminish chromatic aberration of lenses, it was supposed that this defect would make it impossible to build satisfactory large refracting (*i.e.* lens) telescopes. Newton, whose experiments were not precise enough to detect the differences in dispersive power, was of the opinion that powerful telescopes would have to be of the reflecting type.

Example.—*Calculate the focal lengths of two lenses of crown and flint glass which, when placed in contact, will make a converging achromatic pair of focal length* 100 *cm. The refractive indices of crown glass for red and blue light can be taken as* 1·515 *and* 1·523 *respectively, and those for flint glass for the same colours* 1·645 *and* 1·664 *respectively.*

It is first necessary to calculate the dispersive powers of the two glasses from the data. In the absence of any data for the refractive index for the intermediate colour it is usual to take this as being the arithmetic mean of the two refractive indices. Thus the mean refractive index for crown glass is $\frac{1{\cdot}515+1{\cdot}523}{2}$, which is equal to 1·519, so that the dispersive power for crown glass (ω) is given by

$$\omega=\frac{1{\cdot}523-1{\cdot}515}{1{\cdot}519-1}$$
$$=0{\cdot}015$$

Similarly for flint glass

$$\omega'=\frac{1{\cdot}664-1{\cdot}645}{\frac{1}{2}(1{\cdot}664+1{\cdot}645)-1}$$
$$=\frac{0{\cdot}019}{0{\cdot}654}$$
$$=0{\cdot}029$$

It is easier to work in terms of powers rather than focal lengths. The power of the combination is to be 1 dioptre, and if K and K' are the powers of the crown and flint lenses respectively for the intermediate colour, we have the two following conditions from which K and K' can be found :—

$$K + K' = 1 \quad (1)$$

and

$$\omega K = -\omega' K' \quad (2)$$

Eliminating K' from (2) gives

$$\omega K = -\omega'(1 - K)$$

or

$$K = -\frac{\omega'}{\omega - \omega'}$$

$$= -\frac{0{\cdot}029}{0{\cdot}015 - 0{\cdot}029}$$

$$= +2{\cdot}1\mathrm{D}$$

The power of the flint lens is therefore given by

$$K' = (1 - 2{\cdot}1)\mathrm{D}$$

$$= -1{\cdot}1\mathrm{D}$$

Thus the crown lens has a focal length of $\frac{100}{2{\cdot}1}$ cm., *i.e.* 48 cm., and is converging, while the flint lens is diverging and has a focal length of $\frac{100}{1{\cdot}1}$, *i.e.* 91 cm.

Chromatic Aberration Due to a Converging Lens.—It is interesting to consider the extent of the chromatic aberration present in the image formed by an ordinary uncorrected converging lens. We have already discussed the defect when the real image is formed on a screen (page 825), and we now consider the viewing of the image direct with the eye.

In Fig. 635 a converging lens subject to chromatic aberration forms an image of OO′ which is emitting white light. We know that the foot of the

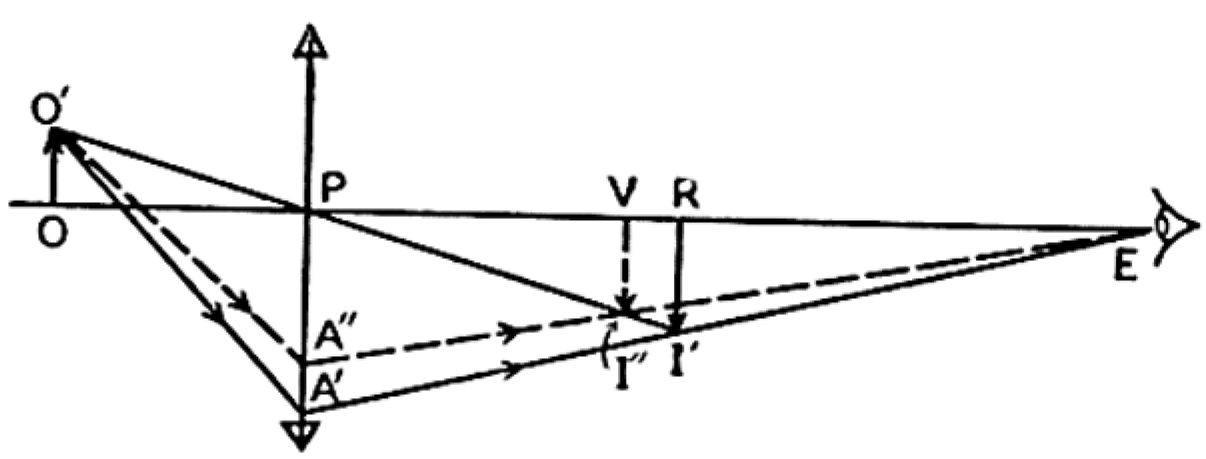


FIG. 635

image will always be on the principal axis and that the heads of the various coloured images will always lie on the auxiliary axis O′P produced, irrespective of how the focal length varies with colour. The focal length is shorter for violet than for red light, so that the relative positions of the violet and red images VI″ and RI′ will be as shown in the drawing. Let an eye E be placed on or near the principal axis beyond the images so as

to see them. The path of the rays by which the eye receives light from I′ and I″ will be (drawing backwards from E) EI′A′O′ and EI″A″O′. Evidently the red image I′ of the top edge of the object is seen without the addition of other colours, but there will be a superposition of colours for images between I′ and I″ until at I″ light from all the images from red up to violet is superimposed on the violet image I″. Chromatic aberration therefore has the effect of colouring the edge of the image red.

We now consider a virtual image due to a converging lens (Fig. 636). Again the heads of the coloured images such as VI″ and RI′ all lie on PO′ produced. If the eye is placed as shown close to the lens so that it can be said to see the images I′ and I″ by receiving the ray O′P, then no colour effects will be seen at the edge of the image because each colour travels by the same path and therefore is superimposed on the rest when it reaches the retina of the eye. In other words, all the coloured images subtend the same angle at the eye and are perfectly superimposed. Thus there is not much chromatic aberration when a converging lens is used to produce a magnified virtual image (see magnifying glass, page 866) provided that the eye is placed close to the centre of the lens. There will be a certain amount of colour, however, if spherical aberration is present, because the amount of this latter defect varies with the colour of the light.

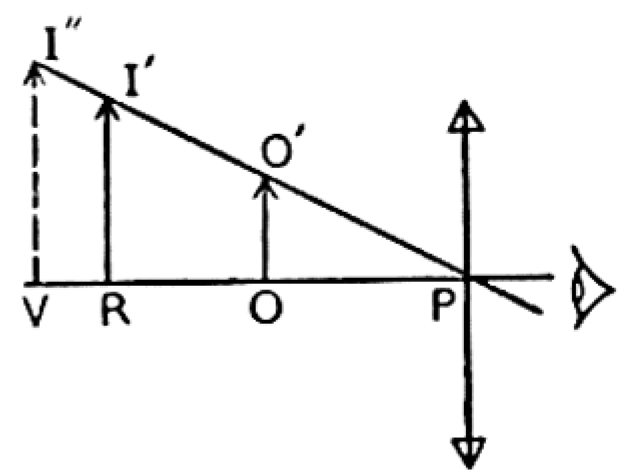


FIG. 636

Achromatism of Two Separated Thin Lenses.*—Another interesting but more complicated case of achromatism occurs when two thin lenses of the same material forming a combination are separated by a distance equal to the arithmetic mean of their focal lengths. This can be proved as follows.

On page 808 it is shown that for a given colour the *reciprocal* of the effective focal length of a system of two lenses separated by a distance a whose focal lengths for that colour are f and f' is equal to

$$\frac{1}{f}+\frac{1}{f'}-\frac{a}{ff'}$$

(The symbols have been changed from those used in equation (15) so as to avoid using the double dashes.)

Both lenses are made of material whose refractive index for the given colour is n, so that

$$\frac{1}{f}=(n-1)B$$

and

$$\frac{1}{f'}=(n-1)B'$$

* See footnote, page 815.

where B is the sum of the reciprocals of the radii of the surfaces of the first lens and B' is the same quantity for the second lens. Therefore the effective power of the system (*i.e.* the reciprocal of its focal length) for the given colour can be written

$$(n-1)(B+B')-a(n-1)^2BB' \quad . \quad . \quad . \quad (7)$$

Now suppose that the colour changes so that the refractive index changes from n to $(n+\Delta n)$. The power of the system then becomes

$$(n-1+\Delta n)(B+B')-a(n-1+\Delta n)^2BB'$$

or

$$(n-1+\Delta n)(B+B')-a\{(n-1)^2+2(n-1)\Delta n+(\Delta n)^2\}BB'$$

If Δn is considered to be only a small change of n, its square can be neglected in comparison with the other terms in the curly bracket, and the last expression for the changed power of the system can be written

$$(n-1)(B+B')+(B+B')\Delta n-a(n-1)^2BB'-2aBB'(n-1)\Delta n$$

The first and third terms in this expression together represent the original power of the system (expression (7)), and the other two terms together represent the change in the power when n changes by Δn. If the system is to be achromatic this change must be zero. Therefore the condition for achromatism is

$$(B+B')\Delta n-2aBB'(n-1)\Delta n=0 \quad . \quad . \quad . \quad (8)$$

Dividing through by Δn and rearranging gives

$$\begin{aligned} 2a &= \frac{B+B'}{(n-1)BB'} \\ &= \frac{1}{(n-1)B'}+\frac{1}{(n-1)B} \\ &= f'+f \end{aligned}$$

Therefore for achromatism

$$a=\frac{f'+f}{2}$$

Students who are familiar with calculus will, of course, be able to arrive at equation (8) by differentiating the power of the system (expression (7)) with respect to n and equating it to zero. The intervening steps set out above are simply differentiation performed from first principles.

The fact that the power of the system is independent of colour means, of course, that the focal length is similarly independent, but this does not mean that incident parallel rays of all colours are brought to the

same principal focus. The fact is that this focus moves as the colour of the light changes, while the principal plane from which the focal length is measured moves in the same way, the focal length remaining constant.

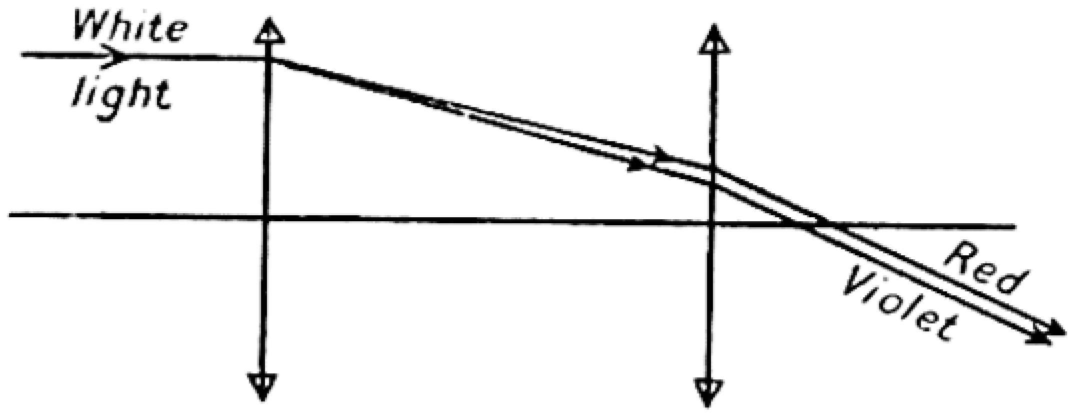


FIG. 637

Fig. 637 is a ray diagram showing that the two coloured rays emerge parallel to each other from the system but are separated. Each undergoes the same total deviation. On entering an eye these two rays (and all the intermediate rays) would be combined on the retina so as to form a white image.

EXAMPLES XLVI

1. Explain the terms *dispersion, dispersive power*.
Explain and illustrate by diagrams the action of (*a*) a direct vision spectroscope, (*b*) an achromatic combination of two lenses in contact.
The refractive index of the glass of a converging lens is 1·64 for red and 1·66 for blue light. If the focal length of the lens is 20 cm. for red light, what will be its value for blue light? (L.I.)

2. Describe the two chief defects in the image formed by a single lens. Show how these defects are corrected in the case of either a Huygens' or a Ramsden's microscope eyepiece. (L.Med.)

3. Find the condition that the focal lengths of a combination of two thin lenses in contact should be the same for red and blue light.
A convex lens of 20 cm. focal length made of glass of dispersive power 0·0150 is placed in contact with a second lens made of glass of dispersive power 0·0195. Find the focal length of the second lens and of the combination, supposing it to be achromatic. (L.I.)

Chapter XLVII

MEASUREMENTS ON LENSES

1. GENERAL REMARKS

In this chapter we shall repeat the procedure adopted in the corresponding section on mirrors and give little more than the theory of the various experiments. The practical details must be learnt in a laboratory.

One important point to be borne in mind is that when real images are formed, chromatic aberration should be avoided by the use of monochromatic light where possible. A sodium flame or lamp is useful for this purpose. The value of focal length or refractive index which is then determined is the one appropriate to the particular light used.

Spherical aberration can be considerably reduced by placing against the lens a card or paper with a central hole of a centimetre or so diameter. The other defects (astigmatism, coma, curvature and distortion) should be kept as small as possible by using small objects. In this way the obliquity of the rays to the principal axis is minimized.

As regards apparatus, an optical bench (page 729) is the ideal, but reasonably satisfactory lens experiments can be performed without this instrument if care is taken in lining up the lenses, etc. and in measuring the distances.

2. CONVERGING LENSES

Distant Object.—The simplest way of estimating the focal length of a converging lens is to cause it to form an image of a suitable distant object on a white screen. In a large room a distant light or window may be used, and in bright daylight clouds with sharp edges are very suitable. It is a good plan to estimate the focal length of a lens in this way before using the more precise methods which are about to be described.

Auto-Collimation.—A suitable object for this—and indeed for many lens experiments—is a hole of a centimetre or so diameter cut in a sheet of metal or card and covered with a piece of ground glass. The hole is covered with wire gauze or, alternatively, two fine wires are stretched across it so as to cross at its centre. The card surrounding the hole is painted white. A monochromatic source is then placed behind the object so that the light, having been diffused by the ground glass, falls on the lens. The gauze or crosswires act as the object. The white area surrounding the hole provides a screen in cases where the object and image

are formed in the same plane, as in the auto-collimation method described here.

The object is set up on the axis of the lens, and a plane mirror is placed on the other side of the lens, facing the lens and with its plane perpendicular to the axis. In this way the light leaving the lens and striking the mirror is reflected back through the lens, and it can be made to form an image alongside the object, *i.e.* on the white surface surrounding the hole, by suitably adjusting the distance between the object and the lens. The corresponding ray diagram is shown in Fig. 638. It shows the paths of two rays, (1) and (2), from O′ to I′.

The only conditions under which OO′ and II′ can lie in the same plane are those under which AM and BM′ are parallel to each other, so that MC and M′D are also parallel to each other and both the

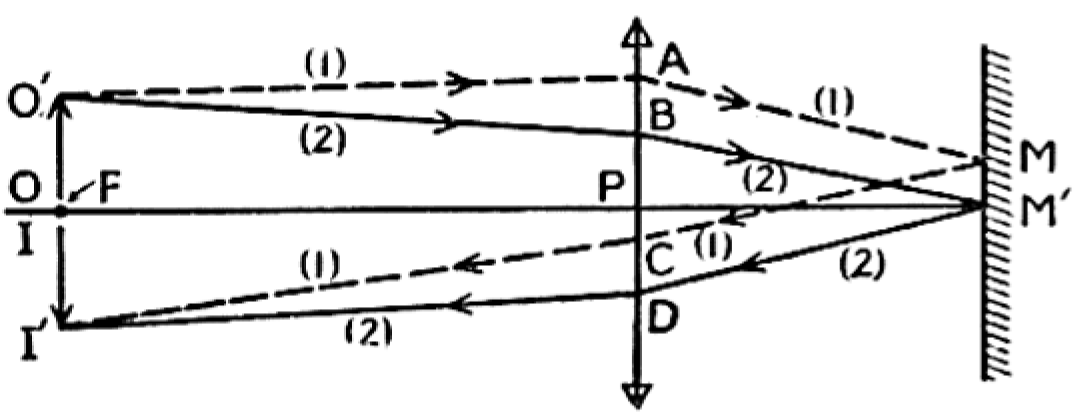


FIG. 638

object and image are in the focal plane of the lens. The focal length is therefore PF. The image is, of course, real, inverted, and the same size as the object.

In performing the experiment it will be found that, as we should expect, the image remains in focus as the mirror is moved towards or away from the lens. It is as well to test whether the correct image is being received by removing or covering the mirror, because it is possible to obtain images formed by reflection at one or other of the faces of the lens itself.

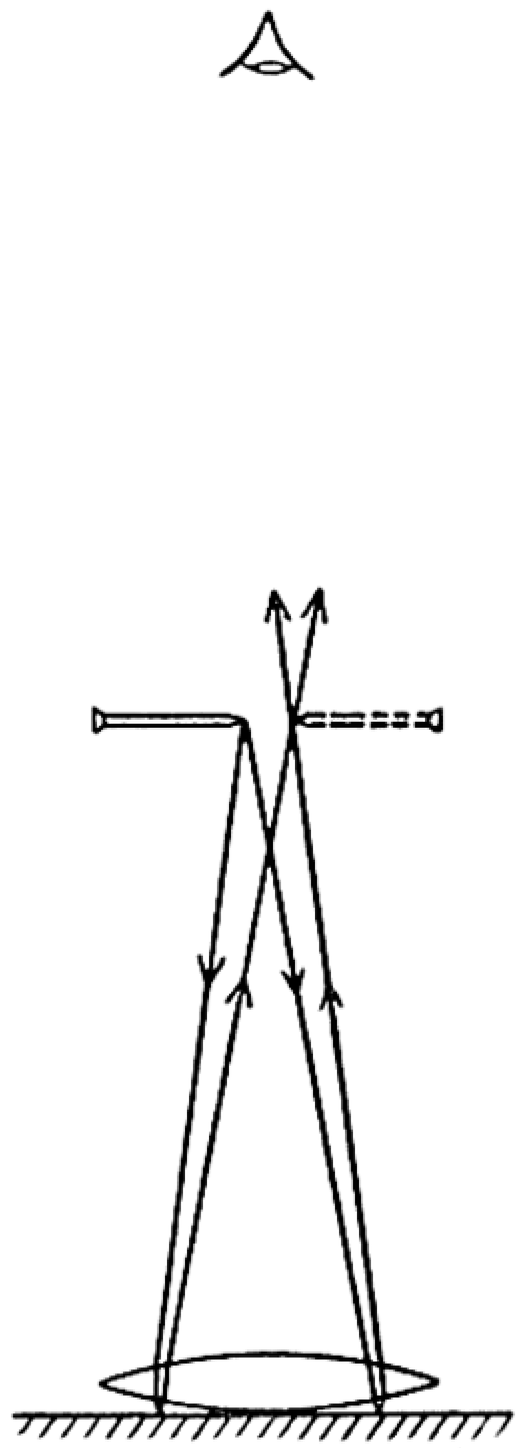


FIG. 639

This method of locating the principal focus of a lens can be used for a combination of lenses and (what is optically the same thing) thick lenses. In this case the position of each principal focus relative to the system can be found by repeating the setting with the lens turned back to front and, in general, the two foci are not symmetrically placed with respect to the system (page 808).

The experiment may be done in a simple way with a thin lens by laying the lens on the horizontal plane mirror (Fig. 639) and holding a large pin horizontally above it with its point vertically above the centre of the lens. The eye then looks down from above and the pin is raised or lowered and moved slightly sideways if necessary, until its real image

is in the position shown in the diagram, *i.e.* in the same horizontal plane as the pin and with its point almost touching that of the pin. This setting can be judged with some accuracy by moving the eye from side to side so as to test for parallax between the point and its image. The distance from the pin-point to the centre of the lens is then the focal length.

Newton's Formula.—In measuring f for a thin lens by the auto-collimation method we should normally measure from the object to the centre of the lens, and the thickness of the latter would not introduce much error especially if f is large. With a thick lens or a combination, however, direct determination of f is impossible at this stage, because we do not know the positions of the corresponding principal planes. When once the principal foci have been *located*, however (*i.e.* their positions determined with respect to some fixed point or points on the lens, such as the centre of the nearest refracting surface), we can then set up an object at a given distance x_1 from one focus and determine the distance x_2 of its image from the other focus. Newton's formula (page 800) then gives

$$f^2 = x_1 x_2$$

so that f can be found. Knowing the positions of the principal foci and the distance of each (f) from its corresponding principal plane, the positions of these two planes can be found.

Conjugate Foci.—Measurements of simultaneous values of u and v for a real object and its real image can, of course, be used for the calculation of f for a thin lens. This method has the disadvantage that the distances should, strictly speaking, be measured to the principal points of the lens and, unless the lens is quite thin, errors may be introduced by the commonly adopted practice of measuring to the nearest surface of the lens or to its geometrical centre. With ordinary lenses, however, such errors are not usually very serious.

Graphical methods may be used to find f from a set of simultaneous values of u and v, for example by plotting the straight line graph of $\frac{1}{v}$ against $\frac{1}{u}$, in which case the intercept on both the axes is $\frac{1}{f}$.

Shift Method.—This method does not require measurements to be made to the lens itself. It can be explained by reference to Fig. 640. When an object is at O and the centre of the lens is at C_1, a real image I is formed provided that OC_1 is numerically greater than the focal length of the lens. For this case we have $u = OC_1$, $v = IC_1$, so that

$$\frac{1}{IC_1} + \frac{1}{OC_1} = \frac{1}{f} \qquad . \quad . \quad . \quad . \quad (1)$$

Now suppose that the object is placed at I. Since the light rays are reversible it is clear that the image would then be formed at O, so that

the distance of the object from the lens would then be IC_1 and that of the image OC_1, and of course the above equation would again be true. This state of affairs, however, can also be brought about by leaving O where it is and moving the lens to some other position C_2 such that $OC_2 = IC_1$ and, by the above argument, the image will again be at I, so that $IC_2 = OC_1$. Thus if a screen remains fixed at I while the object is fixed at O and the lens is moved between them, there are two positions of the lens for which a clear image is formed on the screen, the positions being specified by the above relations. Let the shift of the lens, *i.e.* the distance C_1C_2, be d and the distance from object to screen (OI) be D. Then

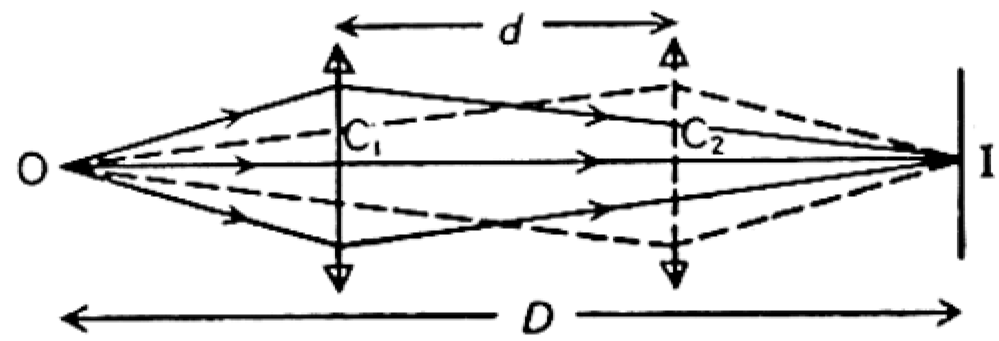


FIG. 640

$$OC_2 - OC_1 = IC_1 - OC_1 = d$$

and

$$IC_1 + OC_1 = D$$

Therefore by adding these last two equations we obtain

$$IC_1 = \frac{D+d}{2}$$

and by subtracting the first from the second

$$OC_1 = \frac{D-d}{2}$$

We now substitute these values for IC_1 and OC_1 in equation (1) and obtain

$$\frac{2}{D+d} + \frac{2}{D-d} = \frac{1}{f}$$

or

$$\frac{4D}{D^2 - d^2} = \frac{1}{f}$$

so that

$$f = \frac{D^2 - d^2}{4D} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Thus the focal length can be determined by observing D and d.

The student should be easily able to see that the magnification is greater than unity for the lens position C_1 and less than unity for the position C_2. Furthermore, each magnification is the reciprocal of the other, and the geometric mean of the two image sizes is equal to the object size. Equation (2) can be written

$$D = 4f + \frac{d^2}{D}$$

and D, d and f are all positive quantities. Therefore, since $\frac{d^2}{D}$ cannot be negative, the minimum value of the right-hand side of this equation is $4f$, which occurs when d is zero. It is evident from the last equation that, for a given lens, as D decreases so must d; that is to say, the two positions of the lens come together until they coincide when d is zero and D is equal to $4f$. If the object and screen are brought still closer together no position of the lens between them will give a real image.

When beginning any experiment involving the formation of a real image by a converging lens, it is useful to remember that this will not be possible unless the distance between the screen and object exceeds $4f$.

In this experiment any convenient mark on the lens holder may be used for measuring the shift of the lens.

Magnification Methods.—The direct measurement of magnification can be used in the determination of the focal length of a converging lens in a way similar to that used for a concave mirror (page 732). We have the equations

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

and

$$-m=\frac{v}{u}$$

Multiplying the first equation by v and combining it with the second we obtain

$$1-m=\frac{v}{f}$$

or

$$v=(-m)f+f. \qquad (3)$$

which shows that a graph of v against $(-m)$ is a straight line with a slope equal to f, and the intercept on the v axis is also equal to f. The intercept on the $(-m)$ axis is equal to -1. It is *not* intended, of course, that m shall be determined from measurements of u and v, but that its numerical value shall be found in any particular case by dividing a linear dimension of the image by the corresponding dimension of the object. For this purpose a suitable object must be chosen such as two narrow parallel slits in a screen or two thin parallel wires or pencil lines on a ground-glass screen. The distance between the real images of the slits, wires, etc. may be measured by forming the image on a screen on which is stuck squared paper or a paper scale, or else by using dividers. A series of values of v and of m found in this way for real images can be plotted and f found from both the slope and intercept of the graph, which is the same as the corresponding graph for a concave mirror (Fig. 542, page 733).

The determination of v involves measurement from the screen to the

lens, and, for reasons already mentioned, a systematic error may be made if the lens is not very thin. This difficulty may be overcome as follows. Suppose that we measure from the screen to *any* convenient point on the lens holder. For the sake of argument (but not necessarily) suppose that this point lies on the image side of the lens and is at a distance s from the principal point to which v should be measured. The distance of the image from the measuring point is then $(v-s)$, and this is the distance which is measured. Let it be V. Then $v=V+s$, and substituting this in equation (3) we obtain

$$V+s=(-m)f+f$$

or

$$V=(-m)f+(f-s)$$

Thus the slope of a graph of V against $(-m)$ is f. The intercept on the axis of V is not f but $(f-s)$ and cannot be used to obtain another value of f. However, since f is found from the slope, s can be determined. Thick lenses and combinations can evidently be treated in this way.

There are other ways of using m for the determination of f. For example, if we multiply the lens equation by u we have

$$-\frac{1}{m}+1=\frac{u}{f}$$

or

$$u=\left(-\frac{1}{m}\right)f+f$$

which can be treated in a similar way to equation (3).

3. DIVERGING LENSES

Because they can be located more easily and accurately, it is usual to arrange for real images to be formed in experiments on the determination of the focal lengths of lenses, and in order to do this when the lens is diverging it is necessary to have a virtual object.

Combination with a Converging Lens.—When a diverging lens is placed in contact with a converging lens of greater power, *i.e.* shorter focal length, the combination is converging, and its focal length f can be determined by any of the methods mentioned in the previous section. It is preferable to use a method not involving measurements of distances to the combination itself, because this will have a finite thickness. The focal length of the converging lens (f'') is determined separately, and that of the diverging lens (f') is then calculated from the relation

$$\frac{1}{f}=\frac{1}{f'}+\frac{1}{f''}$$

In practice the focal length of the converging lens should be considerably

shorter than that of the diverging lens, otherwise the combination will have too long a focal length for accurate measurement.

If the only converging lens which is available is less powerful than the diverging lens whose focal length is required, the following method may be used. The converging lens (whose focal length need not be known) is used to form on a screen a real image I′ (Fig. 641) of a luminous object O, and the distance $I'C_1$ of this image from the lens is measured. The

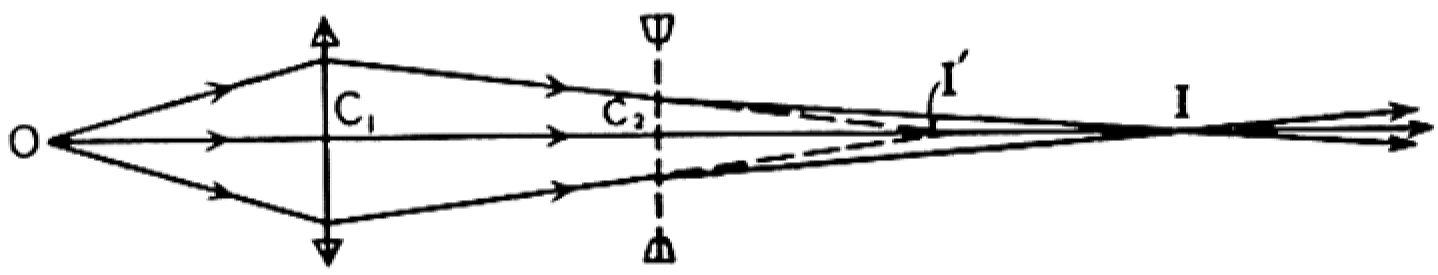


FIG. 641

diverging lens is then inserted between C_1 and I′ and the screen is shifted to obtain the clear image I. Then C_1C_2 and IC_2 are measured. It is necessary to come to the final arrangement by trial and error so as to ensure that the distances to be measured are neither too small nor too large. It is clear that I′ is a virtual object for the diverging lens, and its distance from this lens is found by subtracting C_1C_2 from $I'C_1$. Thus for the diverging lens (focal length f) we have

$$u = -I'C_2 = -(I'C_1 - C_1C_2)$$

$$v = +IC_2$$

which are to be substituted in

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$$

to find f.

Auxiliary Concave Mirror.—The additional converging power necessary for the formation of a real image with a diverging lens can be provided by a concave mirror, as well as by a converging lens. The method is illustrated in Fig. 642. A luminous object is placed in front of a concave mirror in such a position that it coincides with its own real image I′ formed by the mirror. This is simply the location of the centre of curvature as on page 731. The distance I′P is measured. Then the diverging lens is inserted between I′ and the mirror, and it is found to be necessary to move the object away from the lens and mirror to the position O in order to make it again coincide with its real image I formed by the combination of the lens and mirror. The distances OC and CP are measured.

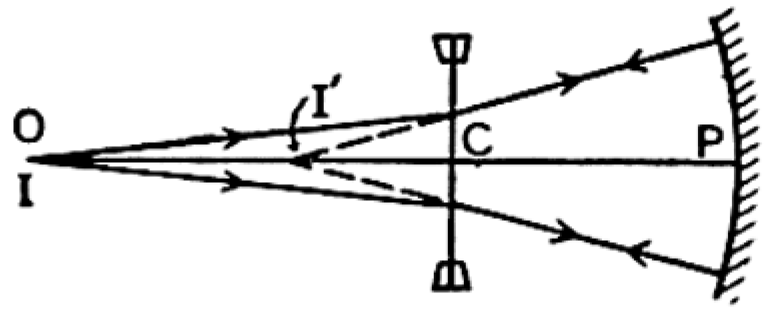


FIG. 642

Since O and I coincide, each incident ray must be reflected back along

its own path by the mirror. Therefore the rays striking the mirror must be travelling along radii, *i.e.* diverging from I′. As regards the lens, therefore, O is a real object and I′ its virtual image. Hence, in the lens equation,

$$u = +\mathrm{OC}$$

and

$$v = -\mathrm{I'C} = -(\mathrm{I'P} - \mathrm{CP})$$

so that f can be found.

Conjugate Foci.—It is, of course, possible to measure directly u and v for a diverging lens and to calculate f from the lens equation. For a real object the image is always virtual, and its position must be found by a parallax method such as that illustrated in Fig. 643, in which the image I of the large pin O is viewed from the opposite side of the lens. The sighting pin P is viewed directly over the top of the lens, and moved about until it is continuous with the image I as the eye is moved from side to side. There are other devices for locating the virtual image, such as the use of a plane mirror similar to the method described for a convex mirror on page 733.

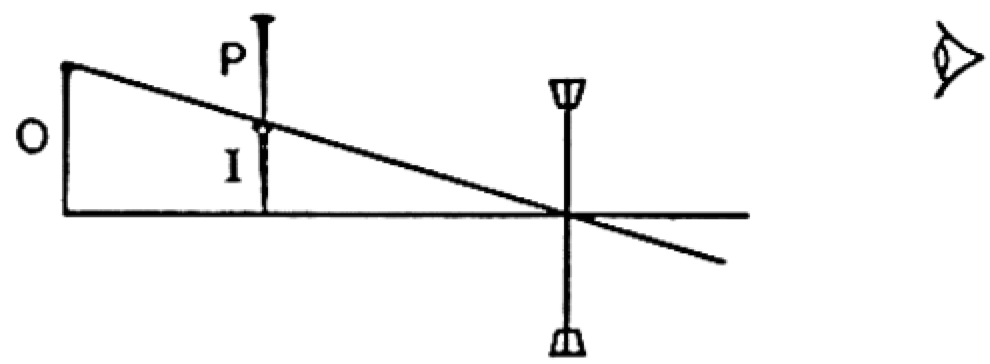


FIG. 643

4. OTHER LENS EXPERIMENTS

Radii of Curvature Optically.—Evidently the radius of curvature of a lens surface which is concave to the air can be determined by any of the methods for concave mirrors, the simplest of which is the location of the centre of curvature as described on page 731. The amount of light reflected from the lens surface is not so great as with a silvered mirror, but it is quite sufficient when a luminous source is used.

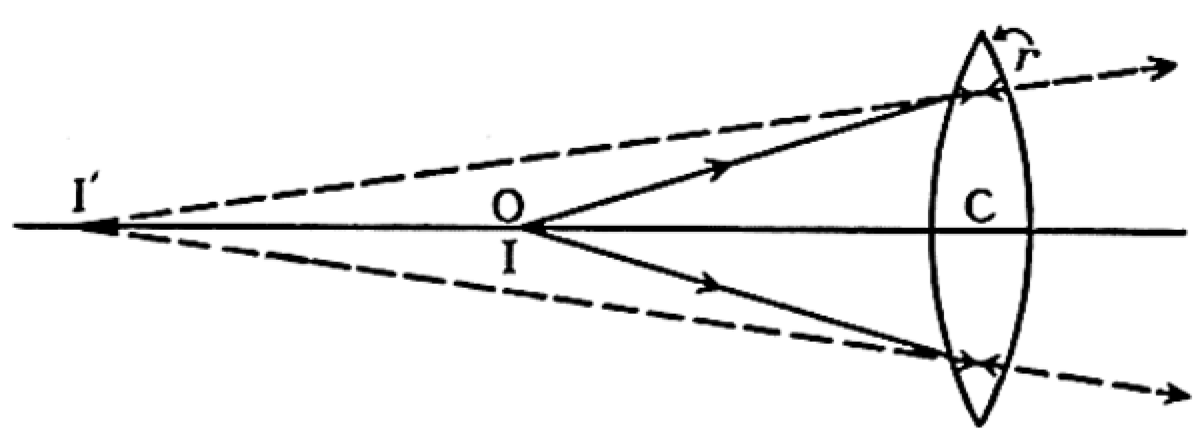


FIG. 644

The radius of a convex surface can be determined by any method applicable to a convex mirror, but it is more usual to use **Boys's method,** which is as follows. A luminous object O (Fig. 644) is moved along the axis of the lens until it coincides with its own image I, formed by reflection at the far surface of the lens (radius r) combined, of course, with two refractions at the near surface.

Since the incident rays are reflected back along their own paths they must strike the far surface of the lens normally, so that when the rays in the glass are produced backwards they must meet at the centre of this surface, which is denoted by I′ in the drawing. Evidently if these rays were not reflected back they would pass through the far surface undeviated (and in fact a large proportion of the light does this), and an eye looking through the lens from the far side would see a virtual image of O at I′. Considering the formation of this image by the lens we have $u=\mathrm{OC}$, and is positive, and $v=r$, and will be found to be negative. Thus if f, the focal length of the lens, has been determined, r can be calculated from the lens equation.

As an alternative to the use of a luminous source, a pin may be used as the object and its correct position found by the method of parallax elimination if the reflection at the lens surface is made stronger. This may be done by floating the lens on mercury with the surface whose radius is required in contact with the mercury. When the focal length and both radii have been determined, the refractive index of the glass of a lens can be calculated from the usual formula. In doing this it must be remembered that, on the convention which we have adopted, the radii of lens surfaces are always considered to be positive quantities. This experiment is often done with a converging meniscus lens. In this case the radius of the convex face is determined by Boys's method, and that of the concave face simply by making the object and its reflected image coincide as with a concave mirror (page 731).

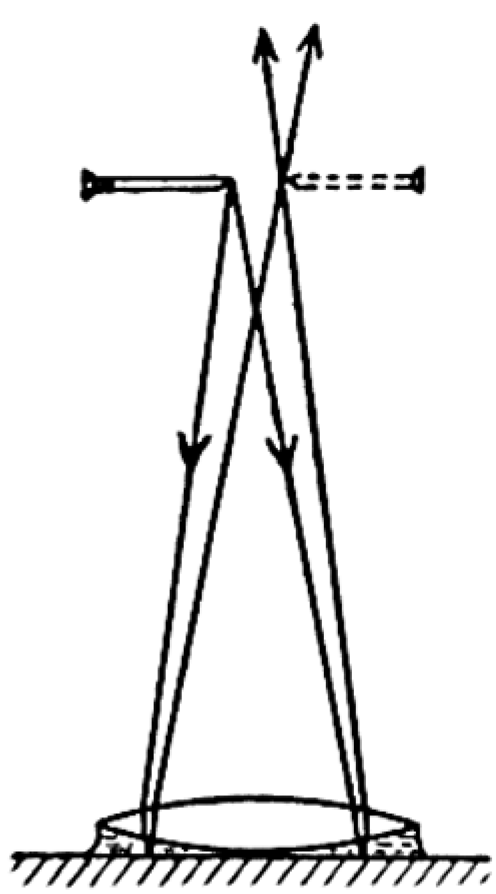


Fig. 645

Refractive Index of a Liquid.—This experiment is a common laboratory exercise. A converging lens (usually double convex) is laid on a horizontal plane mirror, and its focal length (f') is found by making a pin placed above it coincide with its own image (page 835). Next the lens is removed from the mirror, a small quantity of liquid placed on the mirror and the lens replaced. The position in which the pin coincides with its image is now found to be further from the lens (Fig. 645). The height above the lens at which this occurs is numerically equal to the focal length of the combination of the glass lens in contact with what is effectively a plano-concave liquid lens. Let the focal length of the combination be f. It must be taken as positive because the combination is converging. Then the focal length of the liquid lens in air (f'') can be calculated from the equation

$$\frac{1}{f}=\frac{1}{f'}+\frac{1}{f''}$$

and will, of course, be negative.

Lastly, the radius of the glass lens surface which was in contact with the liquid is measured by cleaning the lens and floating it on mercury with the same face downwards, and carrying out Boys's experiment with a pin held above the lens.

Thus the liquid lens has one surface of infinite radius (*i.e.* a plane surface) and a concave surface of radius r, and, if the refractive index of the liquid is n, the focal length of the liquid lens in air (f'') will be given by

$$\frac{1}{f''}=\frac{1-n}{r}$$

since the curved surface would be concave towards the air. Therefore

$$n=\frac{f''-r}{f''}$$

in which, it must be remembered, f'' will be negative and, in accordance with the sign convention we have adopted on page 776, r is, as always, positive.

EXAMPLES XLVII

1. Explain how to find the focal length of a converging lens by the aid of a plane mirror.

A thin lens has one surface convex and the other concave. The radius of the convex surface is 25 cm. and that of the concave surface 50 cm. The refractive index for blue light is 1·520 and the index for red light is 1·510. State, giving your reasons, whether the lens is converging or diverging, and find the difference of the dioptric powers for red light and blue light. (L.Med.)

2. In what circumstances can a prism be said to form an image?

From the consideration of the action of a thin prism on light passing through it (or otherwise), derive the relation connecting the distances of an object and its image from a thin lens with the focal length of the lens.

Draw a diagram to make clear why, when a pin is placed in the focal plane of a converging lens behind which is a plane mirror perpendicular to the axis of the lens, the image of the pin formed in the focal plane is inverted. (L.Med.)

3. A screen is situated one metre from the glowing filament of an electric lamp. A convex lens is placed between them, and two positions of it are found for which a clear image of the filament is formed on the screen. If the distance between these positions is 20 cm., calculate (*a*) the focal length of the lens, (*b*) the ratio of the linear dimensions of the two images. (L.I.)

4. Define *focal length* of a convex lens, and show how its value may be deduced by observing the two possible positions of the lens for which an image is formed on a screen placed at a fixed distance from a luminous object.

A lens is placed in front of, and in contact with, a convex mirror of radius R. If the combination acts as a plane mirror, deduce the nature and focal length of the lens by considering the action of the combination upon an incident parallel beam of light. (L.I.)

5. An object is placed at a distance of 24 cm. from a convex lens, and the image is found to be real and 40 cm. from the lens. A concave lens is then placed between the convex lens and the image, at a distance of 15 cm. from the latter, and it is found that a real image is now formed at a distance of 45 cm. from the concave lens. Find the focal length of each lens. If the object is 2 cm. high, what is the height of the final image? (L.Med.)

6. A convex lens placed 20 cm. from an object forms a real image 30 cm. from the lens. If a concave lens is inserted between the image and the convex lens 18 cm. from the latter, the image moves through a distance of 16 cm. still remaining real. Draw a diagram showing how this last image is formed, and calculate the focal lengths of the two lenses. If the length of the object perpendicular to the axis of the lens is 4 mm., what is the length of each of the two images? (L.Med.)

7. How would you determine the focal length of a diverging lens with the help of (*a*) a converging lens, (*b*) a concave mirror?

A converging lens of focal length 20 cm. forms upon a screen an image 2 cm. long of illuminated slit 1 cm. in length. A diverging lens is then inserted 40 cm. from the converging lens and between it and the screen, the latter being moved until a clear image of the slit is again obtained. If this image is now 3 cm. long, what is the focal length of the diverging lens? (L.I.)

8. Explain, with the aid of suitable diagrams, the formation of images by *reflections* from the front and back surfaces of a bi-convex glass lens. An equi-bi-convex lens of glass whose radii of curvature are 24 cm. is silvered upon one surface. Find the distance from the lens of a luminous point on its axis whose image formed by reflection from the concave silvered surface of the lens coincides with the point. (Assume the refractive index air-glass is $\frac{3}{2}$.) (L.Med.)

9. Describe how you would measure by an optical method the radii of curvature of the surfaces of a double convex lens.

An achromatic telescope objective is made of a plano-concave lens of flint glass, of mean refractive index 1·64 and focal length 156 cm., and a double convex lens of crown glass, of mean refractive index 1·50 and focal length 78 cm., cemented with one face in contact with the concave face of the other lens. Calculate (*a*) the focal length of the complete objective, (*b*) the radius of curvature of each face of the double convex lens. (O.H.S.)

10. Give an account of a method of finding the focal length of a thin concave lens using an auxiliary convex lens which is not placed in contact with it.

A thin equi-convex lens of refractive index 1·50 is placed on a horizontal plane mirror, and a pin fixed 15·0 cm. above the lens is found to coincide in position with its own image. The space between the lens and the mirror is now filled with a liquid, and the distance of the pin above the lens when the image and object coincide is increased to 27·0 cm. Find the refractive index of the liquid. (J.M.B.H.S.)

11. Find the relation between the focal lengths of two thin lenses in contact and the focal length of the combination.

The curved face of a plano-convex lens ($\mu = 1{\cdot}5$) is placed in contact with a plane mirror. An object at 20 cm. distance coincides with the image produced by the lens and reflection by the mirror. A film of liquid is now placed between the lens and the mirror and the coincident object and image are at 100 cm. distance. What is the index of refraction of the liquid? (L.I.)

12. Prove that if two thin lenses are in contact,

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$$

where f_1 and f_2 are the focal lengths of the lenses and f the focal length of the equivalent lens.

An equi-convex lens, radii of surfaces 15 cm., rests on a plane mirror and the intervening space is filled with liquid.

If the focal length of the compound lens is 22·5 cm., find the refractive index of the liquid, that for glass being $\frac{3}{2}$. (L.I.)

13. Define *refractive index.*

Describe in detail how you would determine the refractive index of the glass of a thin converging meniscus lens by optical methods. What is the refractive index if the focal length of the lens is 37·5 cm. and the radii of curvature of its surfaces are 9 cm. and 15 cm.? (L.I.)

14. A thin plano-convex lens is made of glass of refractive index 1·5. When an object is set up 10 cm. from the lens, a virtual image ten times its size is formed. What is (*a*) the focal length of the lens, (*b*) the radius of curvature of the lens surface?

If the lens is floated on mercury with the curved side downwards and a luminous object placed vertically above it, how far must the object be from the lens in order that it may coincide with the image produced by reflection in the curved surface? (L.H.S.)

15. Establish the formula $\frac{1}{f}=\frac{1}{f_1}+\frac{1}{f_2}$, where f_1 and f_2 are the focal lengths of two thin lenses placed in contact and f is the effective focal length of the combination.

An object is placed vertically above the centre of an equi-convex lens which is floating on mercury, and when 8 cm. above the lens it is found to coincide in position with the real image formed by reflection at the lens/mercury surface. The lens is now placed on a drop of water on a horizontal reflecting surface and the object is again placed vertically above its centre. How far from the lens is the object when it coincides in position with its image formed by reflection at the plane mirror near the point of contact with the lens? (n for water $=\frac{4}{3}$; n for glass $=\frac{3}{2}$.) (L.A.)

16. Describe how you would determine the refractive index of a small quantity of a transparent liquid using an equi-convex lens (of unknown optical constants), a plane mirror and pins. A small quantity of mercury in a dish is available if required.

An equi-convex lens rests on a horizontal plane mirror and the space between them is filled with water. An object is placed above the lens-mirror system and coincidence with its image occurs when the object is 15·1 cm. distant from the lens. On substituting oil for the water the object has to be raised 2·8 cm. to restore coincidence of object and image. Determine the refractive index of the oil, assuming that of water to be 1·33; the focal length of the equi-convex lens is 10·0 cm. (L.A.)

Chapter XLVIII

THE EYE AND VISION

1. THE EYE

Structure of the Human Eye.—In the study of elementary Physics we are primarily concerned with the eye as an optical system, so that the following description of its anatomy is in no way detailed.

The eyeball is approximately spherical and of about 1 in. diameter. It is held in the cavity usually called the "socket" by muscles attached to its exterior, and it can be made to swivel in any direction by the action of these muscles.

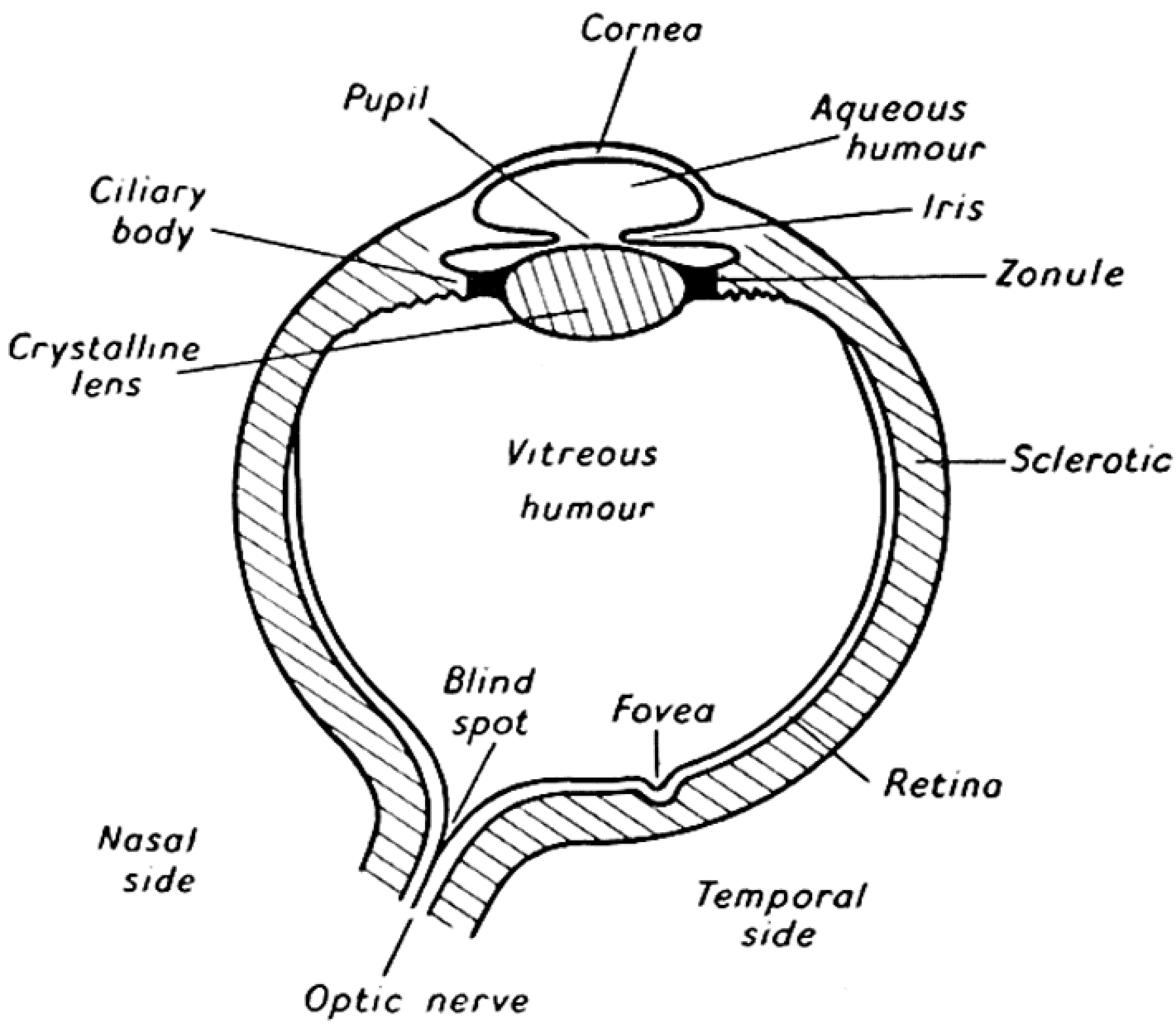


FIG. 646

Fig. 646 is a horizontal section of the right eye. The outside wall covering all but the front part of the eyeball is called the **sclera** or **sclerotic.** The front part is the thin transparent **cornea,** the surface of which is nearly spherical. The chamber immediately behind the cornea contains the **aqueous humour** (refractive index 1·336), and is separated from the larger chamber containing the jelly-like **vitreous humour** (refractive index also 1·336) by the **crystalline lens.** This converging

lens is composed of several layers, of which the inner ones have the greater refractive index. The mean refractive index for the lens as a whole is about 1·40. The lens is partially covered by the diaphragm known as the **iris,** which contains the aperture (the **pupil**) through which the light passes. The pupil ordinarily appears black because no light is reflected out of the eye. Its diameter is controlled by involuntary muscular movements of the iris according to the brightness of the light falling on the eye. The lens is supported by a ring of ligament called the **zonule,** the outer edge of which is attached to the **ciliary body,** consisting of muscle fibres, some of which run radially and some are circular in shape, running round the outer edge of the zonule. The light-sensitive layer, known as the **retina,** covers about five-sixths of the inner wall of the eyeball. Its structure is complicated and its action is not completely understood. It is only necessary to say that the conversion of light into nervous impulse occurs in a layer of the retina containing a very large number of two kinds of minute bodies, which are known as the **rods** and **cones** on account of their shapes. The most sensitive region of the retina is the **fovea** or yellow spot, situated as shown in the diagram. The eye is involuntarily swivelled so that the light from whatever is being looked at falls on the fovea. The light from surrounding objects is simultaneously perceived by the less sensitive parts of the retina.

The nerves along which the stimuli from the various parts of the retina are conducted to the brain all leave the back of the eye at the **optical disc** or **blind spot** as shown in the diagram, the bundle of nerves being known as the **optic nerve.** The fact that the blind spot is insensitive to light can be demonstrated by means of Fig. 647. Using the right eye only,

Fig. 647

look steadily at the cross (its image is then being formed on the fovea) and move the book towards and away from the eye, keeping the line joining the cross and the spot horizontal. At a certain position the spot becomes invisible. Its image is then falling on the blind spot. Rotation of the page in its own plane about the cross will cause the spot to become visible again.

The Optical Action of the Eye. Accommodation.—The greater part of the refraction of light by the optical system of the eye occurs when the light first enters the eye, *i.e.* when it passes into the aqueous humour through the cornea. The effect of this refraction is, of course, to produce considerable convergence, which is added to by the action of the crystalline lens through which the light passes next to form a real inverted image on the retina. The total power of the eye is about +60 dioptres. The

cornea of the normal eye has a radius of curvature of a little less than 8 mm., so that the power of the front surface of the eye separating the aqueous humour (refractive index 1·336) from the air is a little more than $\frac{1000}{8}(1{\cdot}336-1)$ or $+42$D. The power of the crystalline lens itself (about $+19$D in a normal eye focusing a distant object) is considerably smaller than it would be if the lens were situated in air, because it is preceded and followed by the aqueous and vitreous humours whose refractive indices are not much smaller than that of the lens itself.

Human beings have the faculty of altering the power of the refracting system of the eye within a limited range. This is evidently necessary in order that it shall be possible to form clear retinal images, in turn, of objects situated at different distances from the eye. The normal eye can form clear images of very distant objects without there being any sense of strain or muscular effort, and the process of increasing the power of the eye in order to form clear images of nearer objects is called **accommodation.** It is brought about by an increase in the curvature of the front surface of the crystalline lens. The lens is composed of easily deformable (but incompressible) material contained in an outer skin or capsule, and, in the absence of external forces, the lens would become an almost spherical globule. In the unaccommodated eye the lens is stretched by the pull of the zonule ligament on its edge so that its surfaces have their least curvature. It will be remembered that the outer rim of the ligament is attached to the ciliary body which contains radial and circular muscle fibres. In accommodation the circular muscles contract, thus reducing the circumference of the outer edge of the zonule and diminishing the tension in the ligament. The corresponding reduction of tension on the edge of the lens, and therefore in the capsule, allows it to swell at its centre, thus increasing its curvature and optical power. The swelling is almost entirely confined to the front surface of the lens.

The range of distance from the eye within which an object can be clearly focused may be specified by stating the positions of the extremes of the range with respect to the eye. These are known as the **far point** and the **near point** respectively. The distances of these points from the eye are called respectively the longest and the shortest distances of distinct vision. As already explained, the eye is able to cover the range between the two points by accommodation, *i.e.* by changing the power of the crystalline lens. The **amplitude** (or **power) of accommodation** is the change of power which the eye is capable of achieving.

In discussing refraction by the eye it is often convenient to regard the eye as the simple system, known as the "reduced eye," depicted in Fig. 648. This eye is supposed to be filled with a transparent medium, and all the refraction is considered to occur at the front spherical surface, the power

of which we imagine to be variable within the range of accommodation. We shall adopt this method of representing the eye in diagrams.

Suppose that the position of the far point of a particular eye is u_F with respect to the eye, where u_F is subject to the usual real-is-positive sign convention. (We shall see later that it is possible for the far and near points to be the positions of *virtual* objects, so that it is necessary to use a sign convention in order to be quite general.) Let the position of the near point of the eye be represented by u_N, subject to the same sign convention. If an object is moved from the position u_F to the position u_N, the eye will, by definition of u_F and u_N, have to exert the whole of its accommodation in order that the object shall remain in focus. Suppose that, when the object is so moved, the eye remains unaccommodated, but that a lens is placed close to it of such a power that the object can be seen clearly when it is at u_N. This

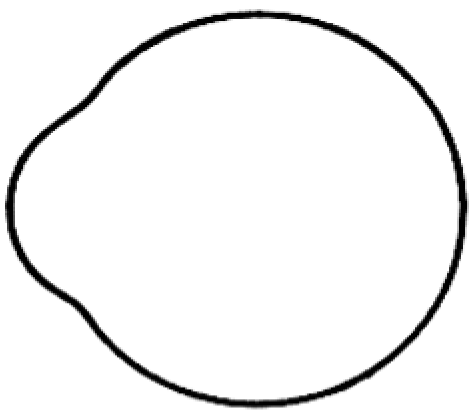


FIG. 648

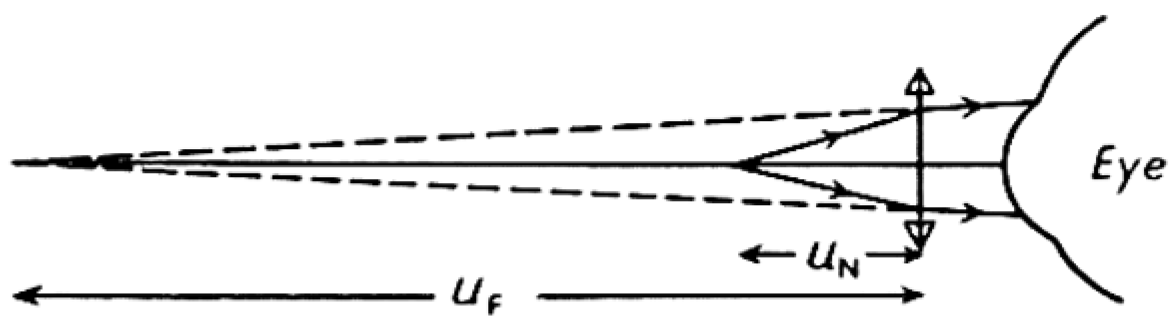


FIG. 649

lens has then had the same effect as the full scope of the accommodation of the eye. Since the eye is actually unaccommodated and it can see the object clearly, it follows that the light actually entering the eye from the auxiliary lens appears to come from an object in the position u_F, whereas the actual position of the object is u_N. In other words, the lens produces an image at u_F when the object is at u_N. Fig. 649 shows a particular case of this in which both the far and near points are situated in front of the eye so that u_N and u_F are both positive since they refer to the positions of real objects. The object is real and the image formed by the lens at u_F is virtual, so that in applying the lens equation to the auxiliary lens we must substitute u_N for u and $-u_F$ for v. By doing this we obtain for the power of the auxiliary lens the expression

$$\frac{1}{u_N}-\frac{1}{u_F} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

which will be in dioptres if the distances are in metres. This expression, therefore, gives the amplitude of accommodation of the particular eye. It is the power of the lens which (if supposed to be placed close to the eye) would cover the observed range of accommodation while the eye itself produces no accommodation. It represents the change of power of which the eye is capable. This formula for the amplitude of accommodation applies to all cases, whether u_F and u_N are positive or negative.

When the far point is at so great a distance from the eye that it can be

said to be at infinity, the eye is regarded as normal, or, to use the technical term, **emmetropic.** It should be realized that the definition of the emmotropic eye is that it is able to see distant objects clearly *when unaccommodated.* A slightly hypermetropic or "long-sighted" eye (page 854) can focus distant objects, but only when a certain amount of accommodation is present. For the emmetropic eye ($u_F = \infty$) it is evident from formula (1) that the amplitude of accommodation in dioptres is simply equal to the reciprocal of u_N in metres. Young people at the age of 20 with normal eyesight have a value of u_N of about 10 cm., thus having an accommodation amplitude of 10D. With increasing age the crystalline lens gradually gets harder and less deformable, so that when the tension in the suspensory ligament and capsule is relaxed in accommodation the lens has less tendency to thicken at the centre, and the amplitude of accommodation is correspondingly less. In other words, the near point moves away in front of the eye with increasing age, and this is a continuous change starting at a very young age. By the time the age of 40 has been reached the near point has moved out to 20 or 25 cm. from the eye, and the corresponding amplitude of accommodation has therefore fallen to 5 or 4 dioptres. At the age of 50 the near point has gone to 40 cm., giving only 2·5 dioptres of accommodation. Beyond this age the further recession of the near point is accompanied by a gradual change in the far point, which becomes virtual, *i.e.* situated behind the eye (as in Fig. 651 (ii)). A similar change begins to occur in the near point after the age of 60, and since both points are then virtual the eye cannot focus real objects at all, no matter where they are placed. It can only focus light within a certain range of convergence. The diminution in the amplitude of accommodation continues meanwhile and no accommodation remains after the age of about 75, that is to say the near and far points are then coincident.

2. DEFECTS OF VISION. SPECTACLES

This section is mainly concerned with a classification of the commoner refraction defects to which the eye is subject, and to the principles underlying their correction by means of spectacles.

The Action of a Spectacle Lens.—We shall begin the discussion by an explanation of the effect which spectacle lenses have on vision. The elementary account which suffices for the student of Physics involves no more than the thin-lens equation. Suppose that a thin spectacle lens of focal length f is placed in front of the eye and so close to it that the distance from any object or image to the eye can be regarded as being the same as the distance of that object or image from the lens. Thus when an object is in a position u with respect to the eye, where u is subject to the real-is-positive sign convention, its position is also u with respect to the spectacle lens. This lens will form an image at a position v with respect to the lens (and to the eye), and v will be subject to the same sign

convention and will be related to u and f by the usual equation

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$$

Therefore, since the light which enters the eye has passed through the spectacle lens, the image formed by this lens at the position v is, effectively, the object for the eye. If this image is virtual, it is situated in front of the spectacle lens and acts as a real object for the eye, while if it is real it acts as a virtual object. Both these cases occur in the following account of the use of spectacles to correct defective vision.

Myopia.—We have seen that as regards distant vision an eye is said to be normal (emmetropic) if it can see distant objects clearly when it is relaxed, *i.e.* unaccommodated. This means that the power of the unaccommodated eye—which is the minimum power the particular eye can have—is such as to bring parallel rays to a clear focus on the retina (Fig. 650 (i)). There are two possible defects in this connection. In the defect which we shall consider first, namely myopia (commonly called "short sight"), the image formed by parallel light falls in front of the retina (Fig. 650 (ii)) when the eye is unaccommodated, so that a distant point object will form a small circle (the "blur circle") instead of a point image on the retina. It is therefore necessary to bring the object towards the eye until it is at the far point in order that it shall be seen clearly (Fig. 650 (iii)). Evidently the structural cause of myopia is either an abnormally large focal power of the refracting system of the eye, or else too large a distance between the front of the eye and the retina. The latter is the more usual cause.

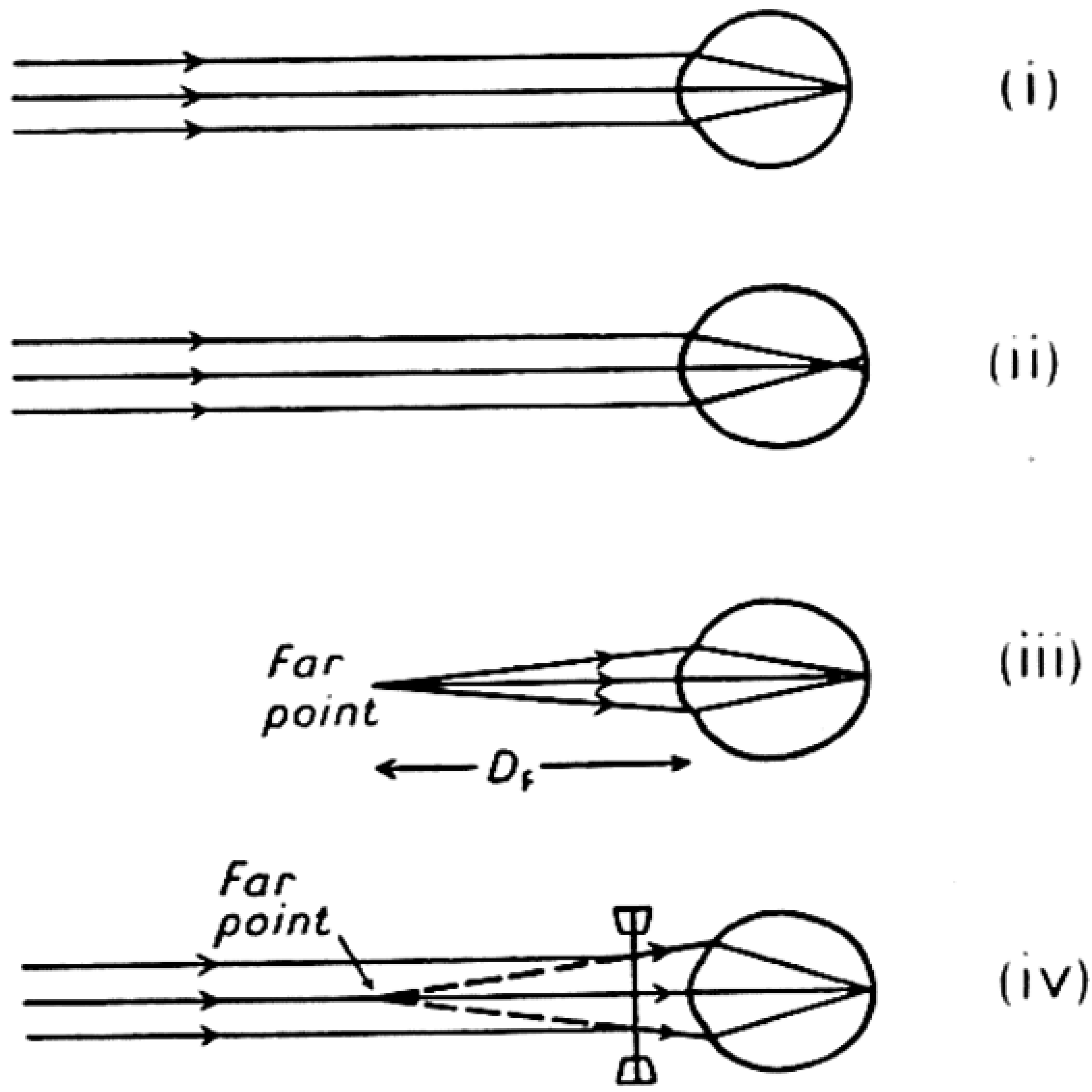


FIG. 650

If the amplitude of accommodation is normal—and there is nothing inherent in myopia to prevent this—then the near point of a myopic eye will be correspondingly nearer to the eye than in normal eyesight. An amplitude of accommodation of 5D in an emmetropic eye would give a near point at $\frac{100}{5}$, *i.e.* 20 cm. in front of the eye, the far point being at infinity. In a myopic eye with a far point say 50 cm. from the eye but

the same amplitude of accommodation, the position of the near point, u_N is given by expression (1) thus:

$$5=\frac{100}{u_N}-\frac{100}{50}$$

so that

$$u_N=\frac{100}{7}$$

$$=14 \text{ cm., approximately}$$

The fact that in myopia the near point is nearer to the eye than for a normal eye with the same amplitude of accommodation is not, in itself, a disadvantage. In fact, it has the advantage of allowing greater magnification when the person is doing close work.

The correction of myopia by spectacles primarily involves the correction of the far point. The principle is very simple, namely that, when a distant object is viewed, the spectacle lens should produce an image of the object at the actual far point of the eye. We shall denote the distance of this from the eye by D_F. Since the far point is in front of the eye and of the spectacle lens, *i.e.* on the same side of the spectacle lens as the object, the image must be virtual, and, since it is to be nearer to the lens than the object, the lens must be diverging. This fact becomes obvious as soon as a ray diagram showing the action of the spectacles (Fig. 650 (iv)) is drawn. Since the light incident on the spectacle lens from any one point on the distant object consists of parallel rays, it follows that the image formed by the lens is at its principal focus. This is a particular case of the general principle of the action of spectacles described on page 850. Therefore, in order that the image shall be at the far point of the eye, *i.e.* a distance D_F in front of the eye, the focal length of the diverging spectacle lens must be numerically equal to D_F, provided that the distance between the spectacle lens and the eye is negligible. Thus we arrive at the very simple principle that myopia is corrected by means of a diverging lens of focal length numerically equal to the distance of the far point from the eye. Ophthalmologists always specify spectacle lenses in terms of their focal power in dioptres, and, although the sign convention which they use in connection with the lens equations is *not* the real-is-positive convention, yet it does give the sign of the power of a converging lens as positive and that of a diverging lens as negative. The ophthalmologist refers to these lenses as "plus" and "minus" lenses respectively. Thus the power of the lens for correcting myopia is equal to

$$-\frac{100}{\text{(distance of far point in cm.)}}\ \text{dioptres}$$

For example, if the far point is 50 cm. in front of the eye, the appropriate lens has a power of -2D. An eye such as this is said to be suffering from

2 dioptres of myopia. A normal eye, or an eye whose far point has been corrected by spectacles, can be made artificially myopic by placing a plus lens in front of it. A lens of power +2D gives 2 dioptres of myopia to a normal eye, +3D gives 3 dioptres, and so on.

When a spectacle lens is worn which corrects the distant vision of a myopic eye the position of the near point will, of course, be changed. In myopia, both the far and near points of the naked eye refer to real objects placed in front of the eye. Suppose that the near point of the unaided eye is situated at a distance D_N in front of the eye and that with the lens the distance is D_N'. The near point when the spectacle lens is worn will be the position in which an object must be placed with respect to the lens in order to produce a virtual image at a distance D_N in front of the lens.

The principle governing the relation between the two near-point positions has been described on page 850. In the thin-lens equation we have

u = position of near point when spectacles worn

$= +D_N'$, since the object is real

v = position of image formed by spectacles

$= -D_N$, since the image is virtual

f = focal length of spectacle lens

therefore

$$-\frac{1}{D_N}+\frac{1}{D_N'}=\frac{1}{f}$$

Thus D_N' can be calculated. If the far point of the uncorrected eye was at a distance D_F in front of the eye, then f is equal to $-D_F$, as explained previously, and we have

$$\frac{1}{D_N'}=\frac{1}{D_N}-\frac{1}{D_F} \qquad . \quad . \quad . \quad . \quad (2)$$

where D_N and D_F are both positive quantities. It will be noticed that the right-hand side of the last equation is the amplitude of accommodation of the uncorrected eye, while the left-hand side is the amplitude of accommodation of the eye when the lens is worn, since under these conditions the far point is at infinity and the near point is at D_N'. We can therefore establish the equation by using the principle that the amplitude of accommodation is unchanged when any particular lens is worn.

***Example.**—An eye has 2 dioptres of myopia and an amplitude of accommodation of 4D. Where are its far and near points? What lens would correct the myopia and where would the near point be when the lens is worn?*

Since the eye has 2D of myopia, its far point is at a distance $\frac{1}{2}$ metre in front of it. In the expression (1) for amplitude of accommodation on page 849, therefore, u_F is equal to +0·5 metre, the positive sign signifying that the position of the far point refers to a real object, as it does in myopia. The position of the near point

(u_N) can then be found from expression (1), knowing that the amplitude of accommodation is 4D. Thus

$$4 = \frac{1}{u_N} - 2$$

so that

$$\frac{1}{u_N} = 6$$

$$\therefore\ u_N = \tfrac{1}{6} \text{ metre}$$

$$= \tfrac{100}{6} \text{ cm.}$$

This is a positive quantity and therefore refers to a real object, as we should expect in myopia.

To correct the far point to infinity a lens is required which will cause an object at infinity to give a virtual image at the far point (page 852). As already explained, this lens must be diverging, and must have a focal length numerically equal to the distance of the far point from the eye. It is therefore specified by the statement that it is a diverging lens of focal length 50 cm. or that its power is -2D.

The position of the near point when the -2D lens is worn can now be calculated by equation (2), or else the following short argument can be used. The amplitude of accommodation is 4D whether the lens is worn or not. When the lens is worn, the far point is at infinity, so that the near-point position u_N is given by expression (1), namely

$$4 = \frac{1}{u_N} - \frac{1}{\infty}$$

$$\therefore\ u_N = \tfrac{1}{4} \text{ metre}$$

$$= 25 \text{ cm.}$$

The fact that this is positive means that the near point is in front of the eye. In fact the vision has been made normal for a middle-aged person by the use of the -2D lens.

Hypermetropia.—This defect of vision, which is sometimes given the shorter name "hyperopia," is commonly called long-sightedness. It can be regarded as the opposite of myopia. The characteristic of a hypermetropic eye is that when it is unaccommodated, *i.e.* when the ciliary muscles are relaxed, parallel light from an object at infinity is brought to a focus *behind* the retina (Fig. 651 (i)). By far the commonest cause of hypermetropia is an abnormally short distance between the front of the eye and the retina. Practically all infants and young children are hypermetropic for this reason, but as they grow up the eyeball lengthens until the eye is emmetropic soon after adolescence. Should the eyeball not lengthen by the usual and correct amount the subject will remain hypermetropic. On the other hand, if the lengthening proceeds too far the eye becomes myopic. This sometimes occurs after adolescence.

The fact that an eye is hypermetropic does not necessarily prevent it from receiving a clear image of a distant object. The eye in Fig. 651 (i)) is unaccommodated, and it is evident that by increasing the power of the eye itself, *i.e.* by accommodation, the rays striking the retina may be made more convergent and possibly (if the accommodation is sufficient) give a clear image. This is not possible, however, if the extent of the hyper-

metropia—as measured by the additional converging power necessary to make incident parallel light give a clear retinal image—exceeds the available amplitude of accommodation. In any case, it is obvious that when any degree of hypermetropia is present, the near point is farther away than with a normal eye having the same amplitude of accommodation, because some of the accommodation is used in providing the power necessary to view distant objects, which is not the case with the normal eye.

In Fig. 651 (ii) the eye is still unaccommodated and the incident rays are shown as having the necessary convergence to give a clear image on

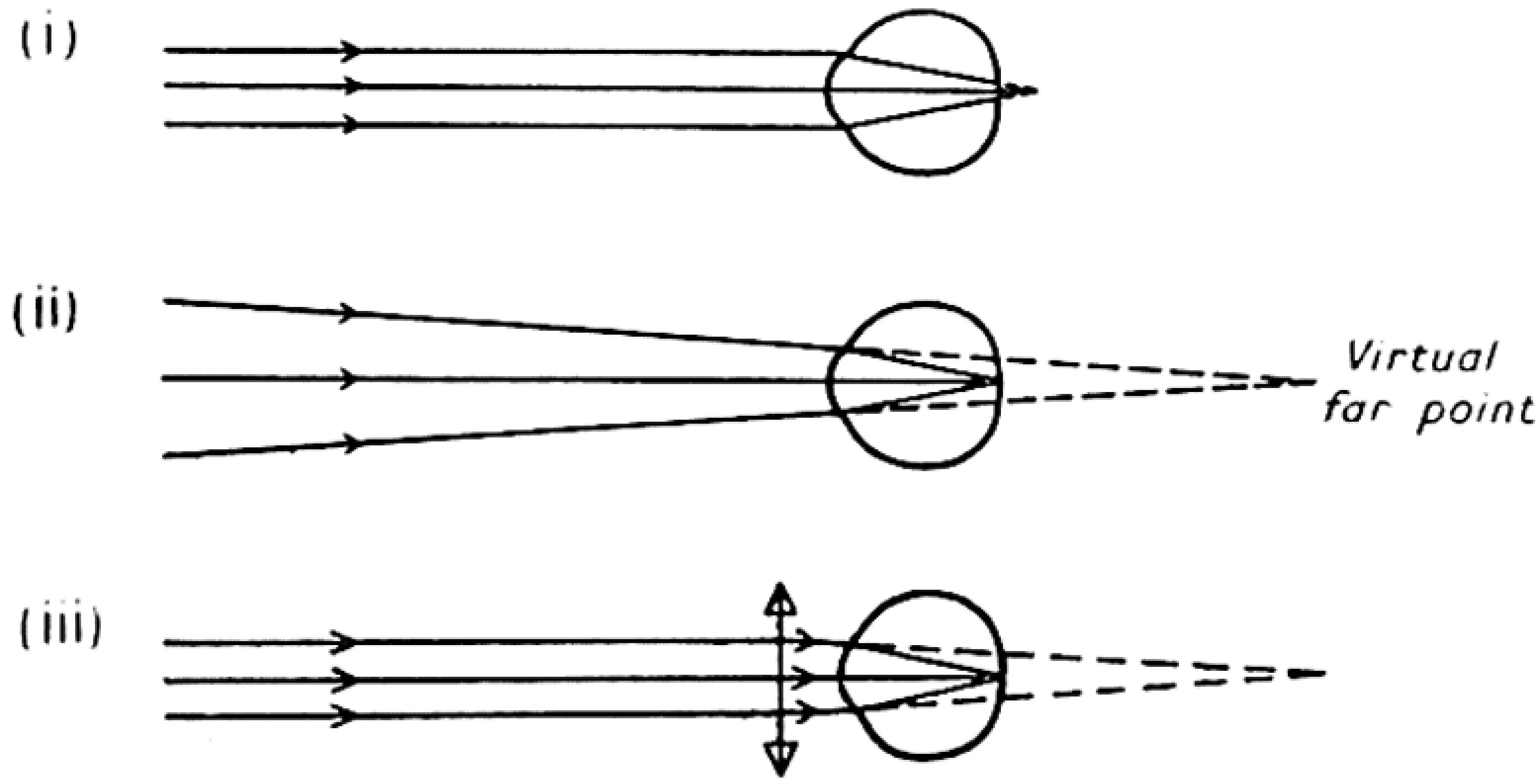


FIG. 651

the retina. The case is analogous to Fig. 650 (iii) (page 851) for the myopic eye, and shows that we can regard the far point of a hypermetropic eye as the position of a virtual object behind the eye. This is the point to which a spectacle lens must converge parallel light from a distant object in order that the object shall be seen clearly when the eye is unaccommodated. It is evident that a converging lens is necessary, and that its focal length must be numerically equal to the distance of the virtual far point behind the eye provided that the distance between the lens and the eye is negligible (Fig. 651 (iii)). The power of this lens is taken as a measure of the hypermetropia of the eye. Thus if an eye has 2 dioptres of hypermetropia its far point can be regarded as situated $\frac{100}{2}$ or 50 cm. behind the eye, and light must be made to converge to this point in order that the unaccommodated eye shall receive a clear image. The near point of the hypermetropic eye may also be behind the eye (but farther from it than the far point) if the degree of hypermetropia exceeds the amplitude of accommodation. In such a case the eye can only focus *converging* rays throughout its whole range of accommodation and it cannot, unaided, see any real object clearly. On the other hand, if the hypermetropia is less than the amplitude of accommodation the eye will have a real near point situated in front of it but, of course, at a greater distance than the near point of a normal eye with the same range of

accommodation. For example, if an eye with an amplitude of accommodation of 5D has 2D of hypermetropia, then 2D of accommodation is required when distant objects are being viewed. This leaves only 3D for accommodation up to the near point, which will therefore be situated at $\frac{100}{3}$ cm. in front of the eye, whereas a normal eye with the same amplitude of accommodation would have a near point at $\frac{100}{5}$ cm. because no accommodation would be needed for distant vision.

It may be noted that, since hypermetropia can be overcome by accommodation, the full extent of the hypermetropia in any particular case cannot be determined unless accommodation is prevented by temporarily paralysing the ciliary muscle. Again, in young people who have plenty of accommodation available a small amount of hypermetropia is no great inconvenience and does not always necessitate the wearing of glasses. With increasing age and decreasing accommodation, however, spectacles become necessary at an earlier age than if the eyes are normal. Finally, the foregoing statement that a hypermetropic eye is made to behave as a normal eye by wearing a converging lens of focal length equal to the distance of the virtual far point behind the eye, must not be taken to mean that the ophthalmologist necessarily prescribes such a lens. The full correction may bring discomfort, especially if glasses have not been worn previously and the patient is therefore accustomed to using accommodation in distance vision. For this and other reasons the full correction is frequently not prescribed and the glasses are often only worn for reading, etc.

The effect of a converging lens on the near point of an eye can be calculated in a way similar to that for a diverging lens, page 853. Suppose that the eye is hypermetropic but that the accommodation is sufficient for real objects to be seen clearly when they are situated beyond a certain point (the near point), which is a distance D_N in front of the eye. Let the focal length of the lens worn be f, and let the near point then be D_N' in front of the eye. This means (Fig. 652) that when a real object is placed at the distance D_N', the spectacle lens forms a virtual image of it at D_N. Thus in the ordinary lens equation $u = D_N'$ and $v = -D_N$, and we have

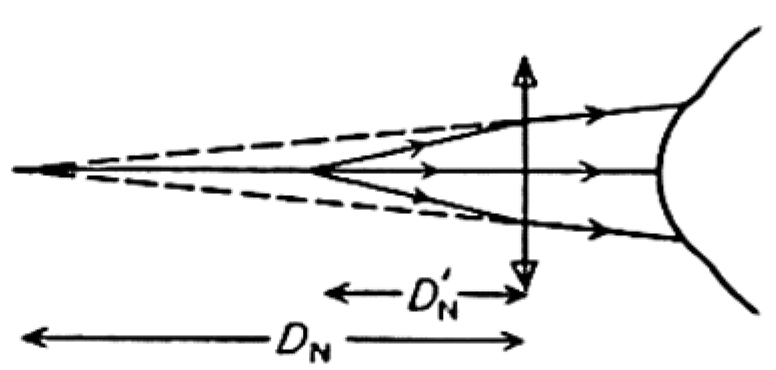


FIG. 652

$$-\frac{1}{D_N}+\frac{1}{D_N'}=\frac{1}{f}$$

Since f is positive D_N' will always be less than D_N, that is to say, the near point is brought nearer by the converging spectacle lens.

In elementary Physics, exercises on hypermetropia frequently consist simply in finding the power of the lens which will correct the near point of a hypermetropic eye to the accepted "normal" value of, say, 25 cm.

In this case the last equation is applicable for finding f when D_N is given as 25 cm. and D_N' as something greater than this.

Presbyopia.—Presbyopia is the name given to the loss of accommodation which occurs as a result of the hardening of the crystalline lens with increasing age. It is likely that this hardening is accompanied by a weakening of the ciliary muscles. The symptom is, of course, increasing inability to see near objects clearly. A converging lens is then required to bring the near point closer to the eye, as explained in connection with Fig. 652 above.

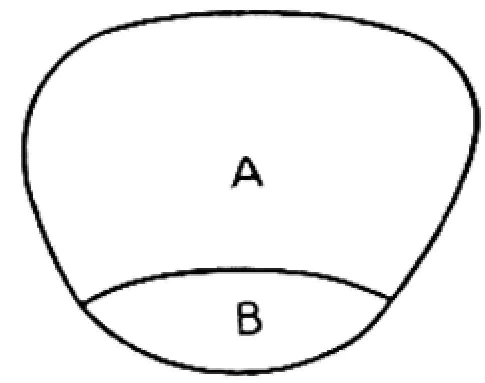


FIG. 653

The average working distance for near work may be taken to be about 30 cm. However if, as the result of age, the whole accommodation of the eye is required at this distance, the person would be judged to have developed presbyopia requiring correction, because it is inadvisable to work at the limit of accommodation. It is evident that for the same rate of development of presbyopia a hypermetrope will need correction for presbyopia sooner than an emmetrope, and an emmotrope sooner than a myope. In fact a myope may never have to be corrected for presbyopia.

Spectacles worn for close work by a presbyope cannot give normal vision for distance because a given lens leaves the amplitude of accommodation unaltered. If the unaided distant vision is normal, then the spectacles for the correction of presbyopia, *i.e.* for increasing the amplitude of accommodation, will be worn only for close work. On the other hand, if the distant vision itself needs correcting for any of the other defects, then separate glasses will be needed for distant and near vision. This inconvenience is often avoided by making bifocal spectacles (Fig. 653), in which the lower portion B of each lens is ground to the necessary power for near vision and the remaining portion A is for distant vision.

Astigmatism.—Astigmatism is the name given to those optical defects which occur when the refracting system of the eye behaves as a combination of a cylindrical and spherical lens (page 811). It is due to the curvature of one of the surfaces of the lens or of the cornea being different in different planes, the shape of the surface being that of a sphere combined with a cylinder. In **simple astigmatism,** rays in one plane are focused on the retina, while in all other planes (and to the largest extent in the plane at right angles to the first) they are focused either in front of the retina (simple myopic astigmatism) or behind it (simple hypermetropic astigmatism). This defect involves an inability to focus all the lines in Fig. 654 with equal clarity at the same time. When one of the lines is clearly in focus the one at right angles is the most out of focus.

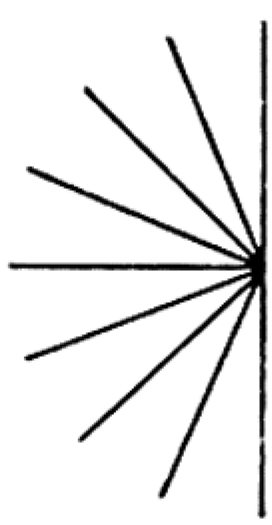


FIG. 654

In **compound myopic astigmatism** the foci corresponding to the two planes are *both* situated in front of the retina, while when both foci

are behind the retina the astigmatism is **compound hypermetropic.** If one focus is in front of the retina and the other is behind, the astigmatism is said to be **mixed.**

The correction of astigmatism requires the use of a cylindrical refracting surface, which in all cases except simple astigmatism must be combined with a spherical surface which will deal with the myopia or hypermetropia. Thus, to take one example, if in compound myopic astigmatism the power of the eye in the plane of maximum power exceeds that in the perpendicular plane of minimum power by, say, 2D, then a diverging cylindrical surface of power -2D with its axis perpendicular to the plane of maximum power will eliminate the astigmatism, leaving ordinary myopia to be corrected by a spherical surface. The same principle applies to hypermetropic and mixed astigmatism. In such cases, therefore, the spectacle lens could have one cylindrical and one spherical surface.

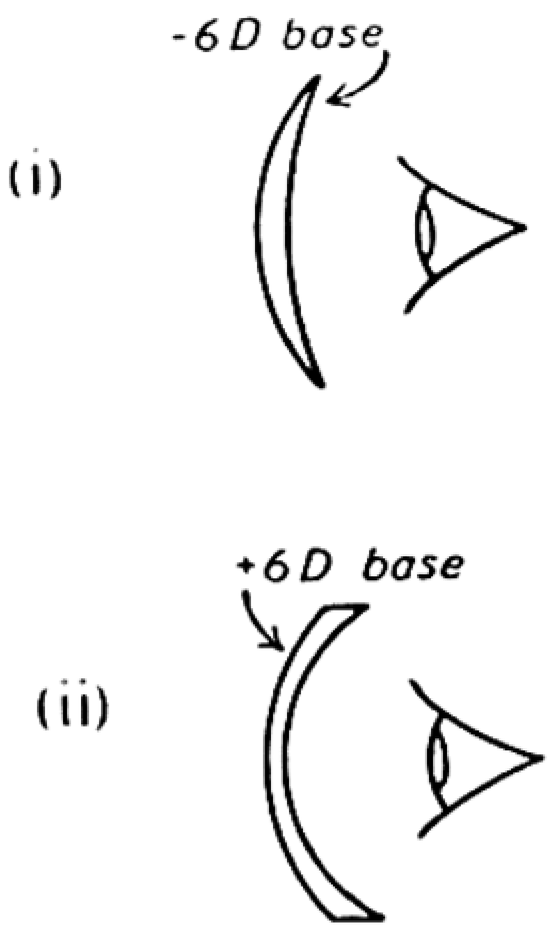


Fig. 655

There are, however, certain points to be taken into consideration when spectacle lenses are being designed. One of the most important of these concerns the astigmatism introduced by the spectacle lens itself when an object is viewed obliquely through the marginal zone of the lens (page 819). To eliminate this disadvantage, spectacle lenses are often made in the meniscus shape (page 786) as in Fig. 655. Such lenses frequently have a 6D base, that is to say, either the back surface (in the case of plus lenses) or the front surface (for minus lenses) is given such a curvature as to contribute a power of 6D to the total power of the lens. If this total power is to be positive, the base curve is placed nearest to the eye (Fig. 655 (i)) and contributes a power of -6D. Thus for a total power of, say, $+3$D the front surface must have a power of $+9$D. The student can easily calculate the radii of curvature of these surfaces on the basis of a value of 1·523 for the refractive index of the crown glass used in spectacles. For a minus lens (Fig. 655 (ii)) the front surface is made the 6D base, and will contribute a power of $+6$D to the lens since it is, of course, made concave towards the eye. Therefore in a lens of power, say, -2D the surface nearer to the eye has a power of -8D. Spectacle lenses made on a 6D base are called **deep meniscus** lenses.

When astigmatism has to be corrected and one of the lens surfaces is the 6D spherical base, it follows that the form of the other surface must be a combination of a cylindrical and a spherical surface. This is the shape of the surface of a circular ring of circular cross-section such as that of an inflated tyre inner tube. The geometrical name for this shape is **toroid,** and a lens of this kind is called a **toric** lens.

Visual Acuity.—The principles and practice of the estimation of refraction errors and the prescribing of spectacles need not be given much attention in a book on Physics, but one or two matters will be mentioned.

For the purpose of trying the effect of lenses of various powers on the vision of a patient, the ophthalmologist—or surgeon, as we shall call him—fits a "trial frame" on the patient. This is an empty spectacle frame of special design and adjustable size. The surgeon also has a set of trial lenses which he can insert in the frame in front of the eye which is being tested. These have a range of powers which enable him, by using them singly or in combination, to produce any desired spherical or cylindrical power between $-10D$ and $+10D$ in steps of 0·25D. When a cylindrical lens is used, the direction of its axis can be adjusted by rotating the lens in the trial frame.

One way of expressing the refraction of an eye is by means of its **visual acuity,** which is arrived at in the following way. Actual measurement on the retinas of normal eyes has shown that the average diameter of the light-sensitive cones is about 0·004 mm. It is reasonable to suppose that in order that a small object (or a large object situated at a great distance) shall be recognizable by the eye, its image on the retina shall cover at least two cones so that its parts can be distinguished. This suggests that, apart from the blurring effect of refraction defects, the smallest retinal image which allows the object to be recognizable measures 0·004 mm. across. This corresponds to an angle of 1 minute subtended by the object at the eye.

In order to test the performance of an actual eye in this respect a set of **Snellen's Test Types** may be used. These consist of rows of square capital letters of different sizes printed on a card, each size subtending an angle of 5 minutes at a distance stated against the row on the card. Each stroke of a particular letter has a width one-fifth of the side of the square which the letter occupies (Fig. 656), so that the width of the stroke subtends the standard angle of 1 minute at the eye at the appropriate distance. Thus any row of letters should be readable by a normal eye placed at the stated distance corresponding to that row. The card is set up at a distance of 6 metres in front of the patient, and each letter of the row labelled 6 m. is then subtending the standard angle. Therefore a normal eye should be able to read this row and the larger ones above it. The patient is asked to read the letters from the top (the biggest) downwards, and the smallest type which he can read is noted. The visual acuity is then expressed as a fraction. The numerator is the distance in metres from the eye to the card (normally 6 metres), and the denominator is the stated distance at which the smallest type which can be read would have to be placed to subtend the standard angle of 5 minutes. It follows that for a normal eye the visual acuity is $\frac{6}{6}$, while if vision is defective (say

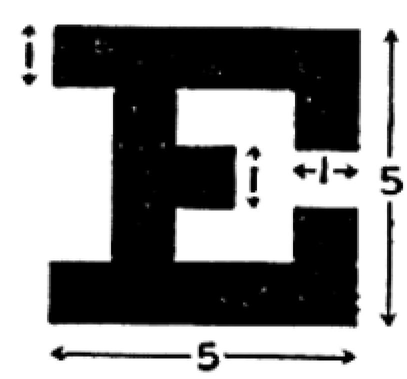


Fig. 656

myopic) so that the smallest letters which can be read would subtend an angle of 5 minutes at 24 metres instead of the 6 metres at which they are actually situated, then the acuity is lower than normal and is $\frac{6}{24}$. It is, of course, possible for visual acuity to be above normal, say $\frac{6}{4}$. When both eyes separately have visual acuities of $\frac{6}{6}$ (or have been made so by spectacle lenses) they reinforce each other when used together and usually give a visual acuity of $\frac{6}{5}$ in binocular vision. This standard is aimed at in prescribing spectacle lenses. It will be realized that the working distance of 6 metres is taken as being equivalent to distant vision, and this is often sufficiently accurate. If not, the difference of $\frac{1}{6}$ dioptre may be taken into account.

The Retinoscope.—We shall mention the principle of only one objective method of sight-testing, namely the retinoscope. In the practice of retinoscopy the ophthalmologist sits facing the patient, his eye O (Fig. 657) being 1 metre from that of the patient, P. A light,

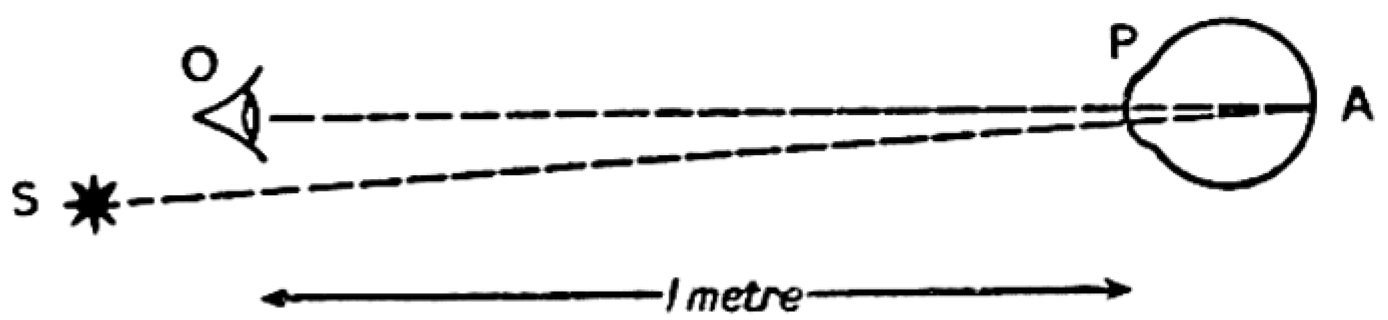


FIG. 657

indicated in the figure by the point source S, is made to shine in the patient's eye. In practice the light is often made to enter the eye by reflection from a plane mirror situated near the surgeon's eye, but the fundamental principle of the retinoscope is unaltered by this. The light from S is converged by the refracting system of the eye P to form a patch of light on the retina. Light from this patch then passes out through the refracting system of the eye, and this forms an image of the patient's illuminated retina at the far point of the patient's eye if, as is arranged, this is unaccommodated. The purpose of the retinoscope is to find the correction necessary to make the far point normal.

Suppose that the patient's eye is myopic to an extent exceeding 1D. The far point is therefore less than 1 metre in front of the eye and is situated between the patient's eye and that of the surgeon. Thus when the latter sees the light coming from the patient's eye he is actually looking at a real image of the illuminated portion of the retina. This is illustrated in Fig. 658, where the position of the image is denoted by I when the source of light is at S. Suppose that the surgeon now moves the source down to S′ so that the patch on the retina moves up to A′. The image then moves down to I′ and the surgeon therefore has the impression that the illuminated patch which he sees in the patient's pupil always moves in the *same* direction as that in which he moves his light source.

Now suppose the eye is hypermetropic. The image is then situated behind the patient's eye (Fig. 659) and is virtual. When the source of light is moved from S to S′, the patch on the retina moves from A to A′ and the image seen by the surgeon moves from I to I′, *i.e.* in the *opposite* direction to that in which the source moves.

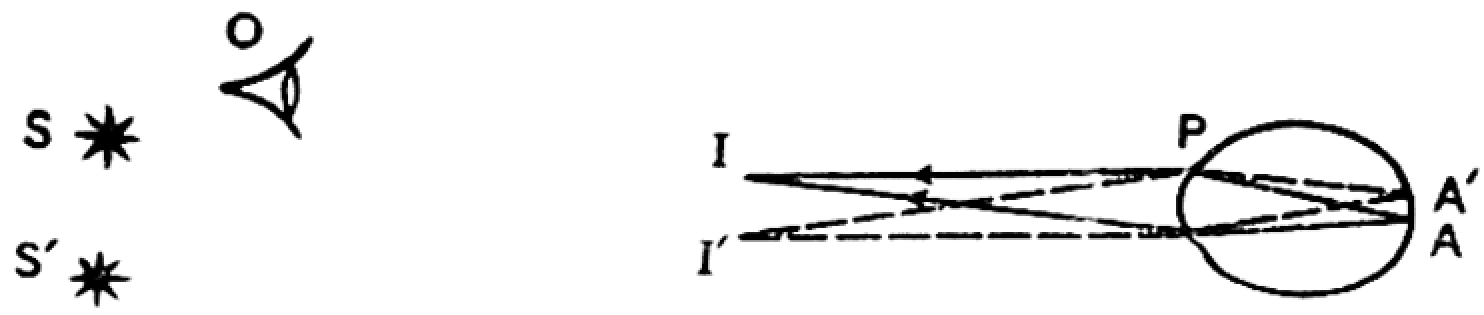


FIG. 658

In practice it is the motion of the edge of the light patch which the surgeon observes, and he is able to judge the nature and extent of the defect and also to neutralize it with a spectacle lens in the following way. Suppose, as we have previously done, that the surgeon's eye is situated 1 metre in front of the patient's. If the patient has 1D of myopia, the image of the illuminated retina will be formed exactly at the surgeon's eye. Therefore, as the source is moved, there will be no apparent movement of the image, only the appearance of light in the patient's pupil when the image is on the surgeon's pupil, and no light when the movement of the

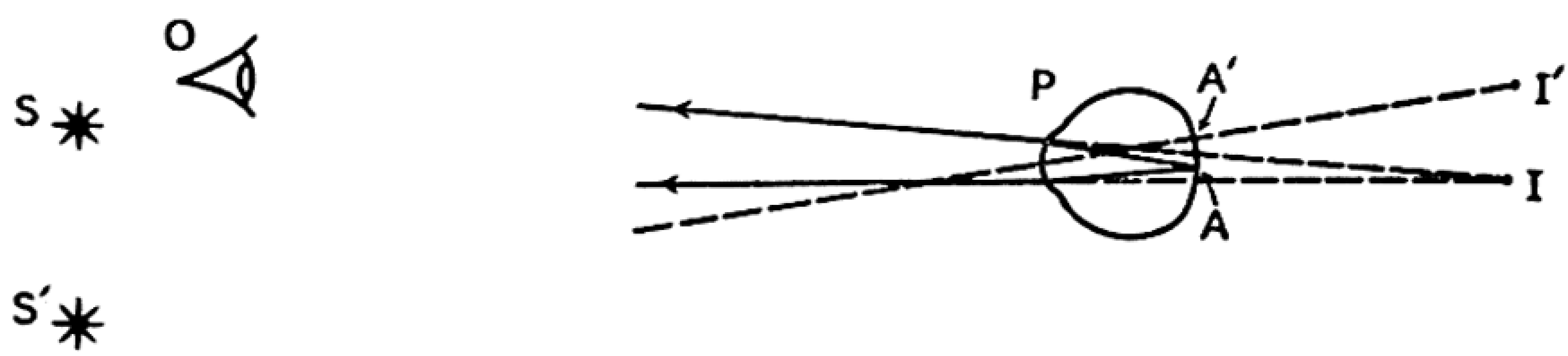


FIG. 659

source causes the image to move off the surgeon's pupil. To look at the matter in another way, the surgeon's eye cannot focus on its own retina the rays forming an image situated on or near its cornea, so there is no possibility of seeing clearly the edge of the patch of light in the patient's eye—the only alternatives are light and darkness. This state of affairs can be recognized with practice, and the surgeon can place spectacle lenses of different power in the trial frame in front of the patient's eye until the condition is reached. He then knows that, with the lens in position, the eye has 1D of myopia, since its far point is situated 1 metre in front of it. To eliminate this myopia he must then add -1D to the lenses already present in the trial frame, or, what is the same thing, subtract $+1$D. The resulting power in the trial frame should then give the eye normal distance vision, and this can be tested by the Snellen chart.

Astigmatism can also be detected and corrected by retinoscopy. The investigation is, of course, more complicated, but the basic principles are the same as those outlined above.

EXAMPLES XLVIII

1. Describe the optical arrangement of the eye, illustrating the description with a labelled diagram.

A person wears bifocal converging spectacles, one surface of each lens being spherical and the other cylindrical. Describe the defects in his vision and explain how the spectacles correct them. (J.M.B.H.S.)

2. Explain how the eye is focused for viewing objects at different distances. Describe and explain the defects of vision known as long sight and short sight, and their correction by the use of spectacles.

Explain the advantages we gain by the use of two eyes instead of one.

A certain person can see clearly objects at distances between 20 cm. and 200 cm. from his eye. What spectacles are required to enable him to see distant objects clearly, and what will be his least distance of distinct vision when he is wearing them? (L.H.S.)

3. Describe how each of the three common cases of defective vision (long sight, short sight and astigmatism) is corrected by the use of suitable spectacles.

A man's near point is 80 cm. from his eye, and his distant vision is normal. What spectacles are needed to enable him to see objects held 25 cm. from his eye, and what will be his greatest distance of distinct vision when he is wearing them? (O.H.S. abridged)

4. Give an account of the eye as an optical instrument. Explain what is meant by the statement that the range of accommodation of a particular eye is 4 dioptres.

A person has a range of accommodation of 2 dioptres and requires spectacle lenses of $+1{\cdot}5$ dioptres to enable him to read a book held at 25 cm. from his eye. Within what range of distances can he see clearly without spectacles? (L.Med.)

5. An eye has an accommodating power of 4 dioptres and its shortest distance of distinct vision is 20 cm. Where is the far point? What spectacles would be necessary for seeing distant objects clearly, and where would the near point be when these are worn? Illustrate your answer by diagrams. (L.I.)

6. A lens combination consisting of a converging spherical lens of focal length 20 cm. in contact with a diverging cylindrical lens of focal length 50 cm. is held 100 cm. in front of a screen. At what distances from the combination can an illuminated slit be placed so as to give a clear image of itself on the screen, and what must be the orientation of the slit with respect to the axes of the cylindrical surfaces in each case?

Relate the observations made with the above arrangement to a common defect of the human eye. (L.Med.)

7. Give a labelled diagram to show the optical arrangements of the eye, and state how it adjusts itself when (*a*) the brightness and (*b*) the distance of the object is varied.

What spectacles would be required by a man whose near and far points are 500 cm. and 150 cm. respectively *behind* his eyes to enable him (i) to read print held 25 cm. in front of his eyes, (ii) to focus distant objects? What is the position of his far point when he is using his "reading" spectacles? (L.I.)

Chapter XLIX

OPTICAL INSTRUMENTS

1. THE SIMPLE MICROSCOPE

In this chapter we shall discuss the principles of certain optical instruments. The class of instrument which we deal with first includes the simple magnifying glass (or simple microscope, as it is often called), the compound microscope and the telescope. The purpose of such devices is to produce magnification, so that small objects (as in microscopes) or distant objects (as in telescopes) can be seen in more detail than with the unaided eye. It is important to realize that the performance of such instruments is not entirely a question of producing a magnified image. This point is made clear in the next paragraph.

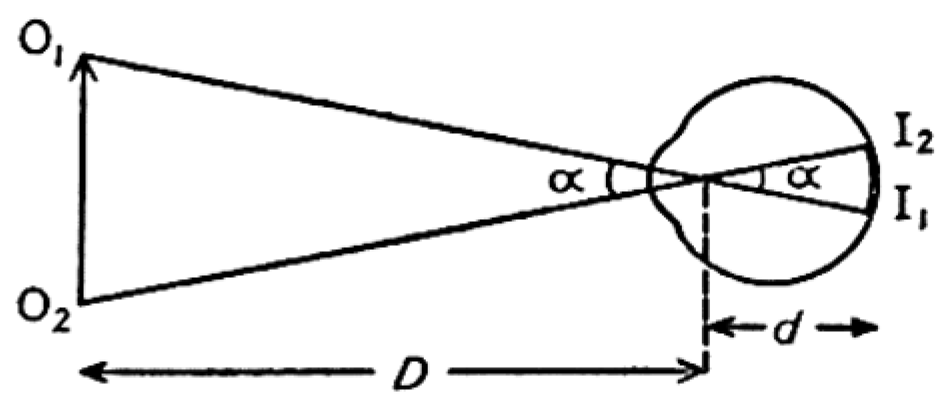


FIG. 660

Angular Magnification.—In Fig. 660, O_1O_2 represents an object situated in front of an eye. It forms an inverted real image I_2I_1 on the retina, and if the optical system of the eye is considered as equivalent to a single thin converging lens, the formation of the image can be represented by the two extreme rays O_1I_1 and O_2I_2 crossing over at the optical centre of the equivalent lens. The angle subtended at the eye by the object is marked α in the diagram, and this is also the angle between the two extreme rays in the eye itself. If the distance of O_1O_2 from the eye is D and the distance of the retina from the centre of the equivalent thin lens is d, then, provided that α is a small angle, we can write

$$\frac{O_1O_2}{D} = \frac{I_2I_1}{d} = \alpha$$

so that

$$I_2I_1 = \alpha d$$

This equation expresses the size of the retinal image in terms of the fixed distance d and the angle which the object subtends at the eye.

When an object is viewed either directly with the unaided eye or through an instrument (in which case it is an image which is viewed), its apparent size is determined by the size of the retinal image I_2I_1, so that the performance of the instrument can be expressed by the relation between the size of I_2I_1 when the instrument is used and its size when the

unaided eye is used. Since d is a constant for any particular eye, the size of the retinal image is directly proportional to the angle subtended at the eye by the object or, when the instrument is in use, by the image formed by the instrument. Clearly, if the size of the object is doubled but it is moved twice as far away from the eye, its size appears unaltered. Suppose that without the instrument the angle subtended at the eye is α and with the instrument it is β. Then the ratio of the corresponding sizes of the retinal image is equal to β/α, and we shall call this the **angular magnification.** It is sometimes called the "magnifying power" of the instrument. Evidently, before a definite value can be assigned to the angular magnification which any particular instrument will give, it is necessary to specify the position in which the object is placed when it is viewed without the instrument (this fixes the value of α), and also the position of the final image due to the instrument when it is used, which fixes β. The positions are just as important as the relative sizes of the object and its image formed by the instrument. We shall now apply this principle to the magnification produced by a single converging lens.

The Simple Microscope.—The action of this device is illustrated in Fig. 661. The lens is placed in front of the object OO′ at a distance less than the focal length, so that a virtual image is formed, which is viewed by an eye placed on the other side of the lens. The base of the object O is supposed for simplicity to be situated on the principal axis of the lens, and the formation of the image is illustrated by locating the image I′ of O′ by means of the usual geometrical construction. The ray O′P passes through the centre of the lens P undeviated, and the ray O′A

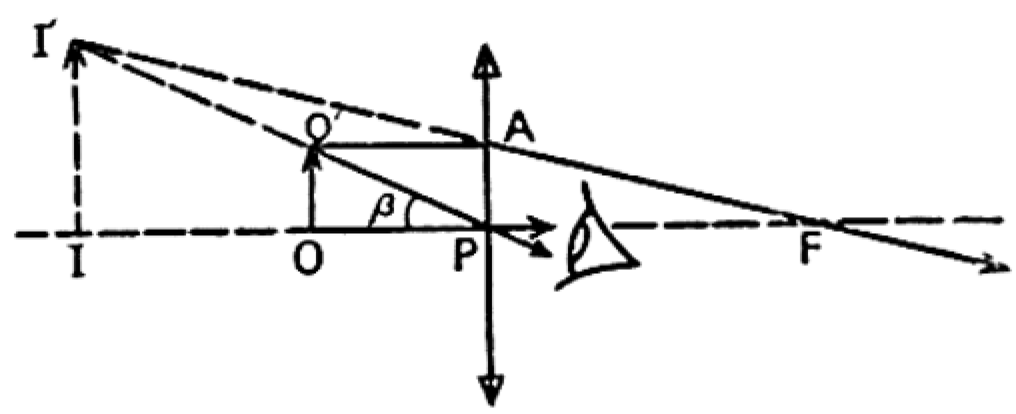


FIG. 661

parallel to the axis is refracted through the second principal focus F. A ray from O along the axis passes through P undeviated. In order to give a simple treatment of the angular magnification, we must suppose that the eye is placed right up against the lens. In this case the angle subtended by the image at the eye is equal to $\widehat{I'PI}$, which is marked β in the drawing.

In order to express the performance of the lens in terms of angular magnification, we have to compare β with the angle which the object subtends at the eye when viewed directly without the lens. It might at first sight be supposed from Fig. 661 that this latter angle is also β, as indeed it would be if the lens were removed and the object and the eye

remained in the positions shown in Fig. 661. In that case the lens would be producing no effective magnification. Although it forms a magnified image, the distance of the image from the eye is such as to make the angle subtended at the eye the same as when the object is viewed direct. Thus, in order to discuss the magnification which we know from experience can be obtained with a single converging lens, it is necessary to specify the position relative to the eye of the object when the lens is not used and of the image when the lens is used.

In discussing vision we have seen that an eye can focus clearly an object situated anywhere between the two limiting points known respectively as the far point and the near point. It is obvious that a given object appears largest (*i.e.* the largest retinal image is produced) when it is situated at the near point. This is the position, therefore, in which a small object would naturally be held for examination by the unaided eye, and it is logical to compare the performance of the magnifying glass with that of the naked eye under these conditions. When the lens is used, the image can be made to occupy any desired position between the near and far points of the eye by moving the object relative to the lens, and, as with the object viewed directly, the position for greatest magnification is the near point. We shall therefore work out the angular magnification by comparing the angle subtended at the eye when the lens is used in such a way that the image is at the near point with the angle subtended when the object is held at the near point and viewed without the lens. If the transverse dimension of the object is represented by the length OO′ and the near point is a distance D_N in front of the eye, the latter angle, which we shall call α, is given by

$$\alpha = \frac{OO'}{D_N}$$

If the image formed by the lens is supposed to be at a distance D_N from the eye, and therefore from the lens, we can write, referring to Fig. 661,

$$u = OP \qquad \text{(real object)}$$

$$v = -IP = -D_N \quad \text{(virtual image)}$$

so that

$$-\frac{1}{D_N} + \frac{1}{OP} = \frac{1}{f}$$

whence

$$\frac{1}{OP} = \frac{f + D_N}{fD_N}$$

The angle β subtended by the image at the eye is given by

$$\beta = \frac{II'}{IP} = \frac{OO'}{OP} = OO' \times \frac{f + D_N}{f.D_N}$$

Therefore the angular magnification (β/α) is given by

$$\frac{\beta}{\alpha}=\left(OO' \times \frac{f+D_N}{f.D_N}\right) \div \frac{OO'}{D_N}$$

$$=\frac{f+D_N}{f}$$

$$=1+\frac{D_N}{f} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This, therefore, is the expression for the angular magnification if the image is at the near point when the lens is used and the object is at the near point when the lens is not used. The arrangement in the two cases is shown in Fig. 662. As we might expect, the magnification depends on the actual position of the near point of the observer as represented by D_N. If this is assumed to have a certain value (*e.g.* 25 cm.), the angular magnification can be worked out in the case of any particular lens of known focal length. It may be noticed that the angular magnification is actually equal to the ratio of the size of the image to that of the object. This is so because the case considered is one in which the comparison is made between the angles subtended by the final image and the object when they are *both* at D_N from the eye.

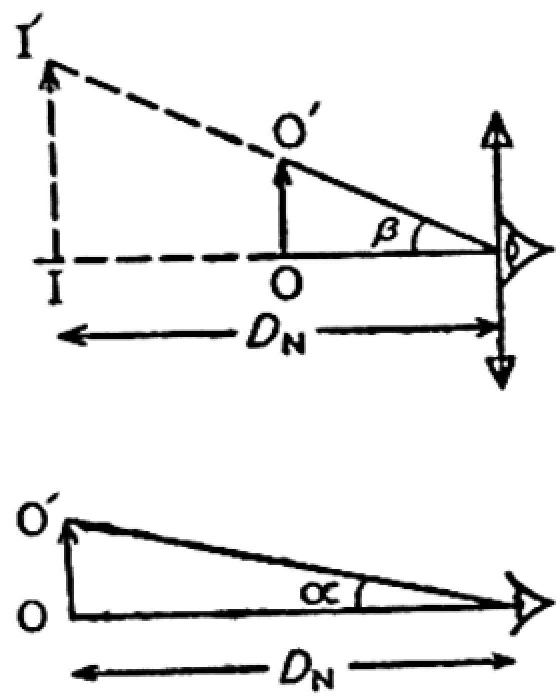


FIG. 662

The angular magnification can be estimated directly by using two objects of the same size, *e.g.* equal divisions on two scales. One object is set up at the distance D_N in front of one eye, while the other is viewed through the lens held close to the other eye. The distance of this object from the lens is adjusted until its image is at the near point and then, by looking with both eyes at once, it is possible to estimate how many times the size of the image exceeds that of the object viewed directly.

The eye is using its full accommodation when it is looking at an object or image situated at its near point. It is sometimes more comfortable, therefore, to have the image further from the eye than D_N when using the lens. In this case the angular magnification is somewhat reduced. In the extreme case in which the image is at infinity so that the eye is completely relaxed (if it is emmetropic), we have in $v=\infty$ the above analysis. It is easy to see that, provided we still suppose that when the lens is not used the object is held at the near point, this leads to the expression

$$\frac{\beta}{\alpha}=\frac{D_N}{f}$$

Thus the angular magnification diminishes by unity when the image moves from the near point to infinity.

The angular magnification obtainable with a simple magnifying glass is usually calculated from the simple formula $\frac{25}{f}$, where f is the numerical value of the focal length in centimetres. Thus a lens with a focal length of 2·5 cm. (about 1 in.) is said to magnify ten times (written $\times 10$). To give large magnification a lens must have a short focal length, *i.e.* highly curved surfaces, and this causes the lens to be thick. Magnifying glasses of this type are used, and so also are various combinations of two or three comparatively thin lenses. Really large magnification, say more than about $\times 25$, can be achieved only by using the principle of the compound microscope.

On page 831 it is shown that chromatic aberration is small in the simple microscope, because the various coloured images formed by the constituents of white light are superimposed when the eye is placed close to the lens.

2. THE COMPOUND MICROSCOPE

The magnification produced by a compound microscope occurs in two distinct stages. The basic principle is to use a short-focus converging lens, called the **objective,** to produce a magnified *real* image of the object, and then to view this image through another converging lens (the **eyepiece**)

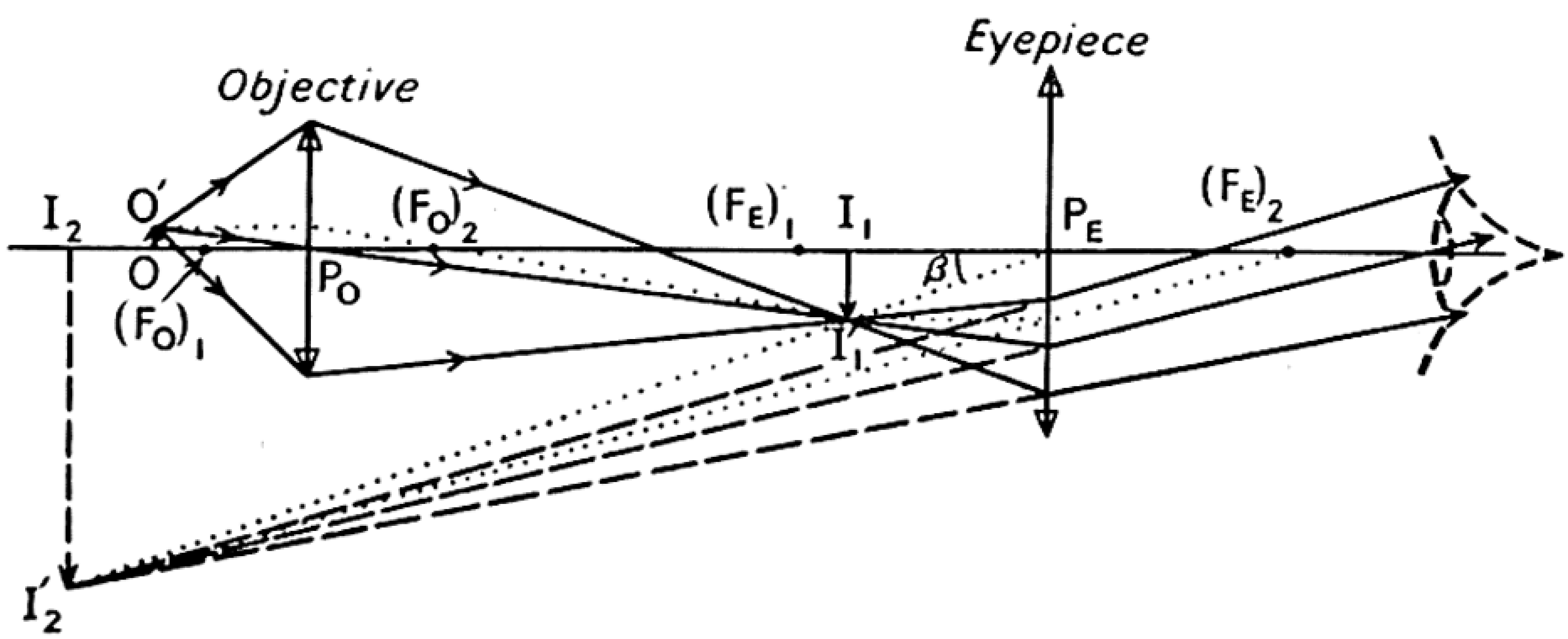


FIG. 663

which acts as a simple microscope. Thus the final angular magnification due to the instrument, as compared with the performance of the unaided eye, is equal to the angular magnification due to the eyepiece multiplied by the linear magnification of the first real image compared with the object.

Fig. 663 is a ray diagram which is made by using the graphical construction first to find the real image I_1I_1' which the objective forms of the object OO', and then to find the virtual image I_2I_2' which the eyepiece produces of I_1I_1', the latter acting as a real object for the eyepiece. For

the purpose of these constructions, which are shown by the dotted lines, the second principal foci $(F_O)_2$ and $(F_E)_2$ of the objective and eyepiece respectively are marked on the diagram. The first principal foci of these lenses are $(F_O)_1$ and $(F_E)_1$ respectively. It should be noticed that in order that the magnification of the first real image shall be considerable, the object must be only a short distance on the far side of $(F_O)_1$ from the objective. The image I_1I_1' must be just within the focal distance of the eyepiece, *i.e.* just to the right of $(F_E)_1$, so as to form a virtual image. The two lenses are usually fixed in the microscope tube, and the position of the final image can be adjusted by the observer by moving the whole microscope tube towards or away from the object, thus altering the position of I_1I_1'.

The diagram (Fig. 663) also shows the path of a pencil of rays from a non-axial point O′ on the object, through the lenses, to the eye. The student should make sure that he is able to draw the path of such a pencil after having found the positions of the images by the construction. It should be noticed that the bounding rays of the pencil continue straight on when they cross over at the point I_1' on the real image.

Angular Magnification Due to Compound Microscope.—If P_O (Fig. 663) is the centre of the objective, it is clear that the linear magnification of the first image is given by

$$\frac{I_1I_1'}{OO'}=\frac{I_1P_O}{OP_O}$$

so that

$$I_1I_1'=OO'\times\frac{I_1P_O}{OP_O} \qquad . \quad . \quad . \quad . \quad . \quad (2)$$

Applying the usual lens equation to the eyepiece (centre P_E) and supposing that the final image is at the near point of the eye, we have

$$u=I_1P_E$$

$$v=-(I_2P_E)=-D_N$$

$$\therefore \quad -\frac{1}{D_N}+\frac{1}{I_1P_E}=\frac{1}{f_E}$$

whence

$$\frac{1}{I_1P_E}=\frac{f_E+D_N}{f_ED_N} \qquad . \quad . \quad . \quad . \quad . \quad (3)$$

To obtain the largest field of view the eye should be placed in the position shown in Fig. 663. In practice the distance between this position and the eyepiece itself is small compared with I_2P_E, and we may therefore state that the angle subtended at the eye by the final image I_2I_2'

is β and is given by

$$\beta = \frac{I_2I_2'}{I_2P_E}$$

$$= \frac{I_1I_1'}{I_1P_E}$$

$$= I_1I_1' \times \frac{f_E + D_N}{f_E D_N} \qquad \text{by equation (3)}$$

$$= OO' \times \frac{I_1P_O}{OP_O} \times \frac{f_E + D_N}{f_E D_N} \qquad \text{by equation (2)}$$

If the object itself were viewed by the unaided eye at a distance D_N it would subtend an angle α given by

$$\alpha = \frac{OO'}{D_N}$$

Therefore the angular magnification due to the compound microscope is given by

$$\frac{\beta}{\alpha} = \frac{I_1P_O}{OP_O} \times \frac{f_E + D_N}{f_E}$$

$$= \frac{I_1P_O}{OP_O}\left(\frac{D_N}{f_E} + 1\right)$$

As was previously stated, this expression is equal to the product of the magnification due to the objective, and the magnifying power of the eyepiece regarded as a simple microscope. If the final image is at infinity instead of at the near point of the observer, so that the distance between I_1I_1' and the eyepiece is equal to the focal length of the latter, the expression becomes

$$= \frac{I_1P_O}{OP_O} \times \frac{D_N}{f_E}$$

which means that the angular magnification is reduced by the number $\frac{I_1P_O}{OP_O}$.

It is evident that in order to obtain large magnification by the compound microscope it is necessary that the first magnification shall be large. It is sometimes as high as 100. With an instrument in which the length of the tube is fixed, the distance of the first real image from the objective (I_1P_O) is fixed, and the magnification due to the objective is inversely proportional to the distance OP_O between the objective and the position in which the object must be placed in order to produce the first real image at the right place. This distance diminishes as the focal length of the objective is decreased. For high magnification, therefore, an objective with a short focal length (high power) is used. When the focal

57

length is very short, say $\frac{1}{4}$ inch or less, as it often is, the object must be situated almost at the principal focus in order to produce the first real image in the correct position in front of the eyepiece, and, as an approximation, we may substitute f_O for the distance OP_O in the expression for the magnification. For the case of the final image at infinity we therefore write

$$\frac{\beta}{\alpha}=\frac{(I_1P_O)\times(D_N)}{f_O f_E}$$

By way of example we can take I_1P_O as about 6 in., D_N as 10 in., f_O as $\frac{1}{4}$ in. and f_E as 1 in. These figures give an angular magnification of $\frac{6\times10}{\frac{1}{4}\times1}$, *i.e.* 240. It will be understood, of course, that we have treated the objective and eyepiece as thin lenses, which in practice they are not, as will be mentioned later.

It is quite possible for a compound microscope to form a *real* magnified image by arranging for the first real image to fall outside instead of inside the focal distance of the eyepiece. This can be done either by withdrawing the eyepiece beyond its usual position, or by increasing the distance of the object from the objective and so moving the first real image away from the eyepiece. A magnified real image is then formed beyond the eyepiece, *i.e.* outside the microscope, and it can be caught on a screen if the illumination is sufficient. Further, the image may be photographed if the screen is replaced by a sensitive plate and extraneous light is excluded by bellows extending from the eyepiece to the plate-holder.

Experimental Estimation of Angular Magnification.—The estimation of the angular magnification due to a compound microscope can be done by a direct-comparison method similar to that for a simple microscope (page 866). A finely-graduated linear scale (S_1, Fig. 664)—a divided millimetre etched on a glass microscope slide is very suitable—is placed on the stage of a microscope. A thin glass plate P is placed immediately in front of the eyepiece and arranged at 45° to the axis of the microscope. The purpose of this is to allow the eye, at the same time as it sees the magnified image of S_1 through the microscope, to see a scale S_2 graduated in the same units as S_1 and placed in such a way that light from it travels the standard distance of 25 cm. before reaching the eye. The microscope is raised or lowered until the image of S_1 shows no parallax with S_2. They are then both at the standard near point, and a comparison of the apparent sizes of equal divisions can be made. This is equal to the angular magnification.

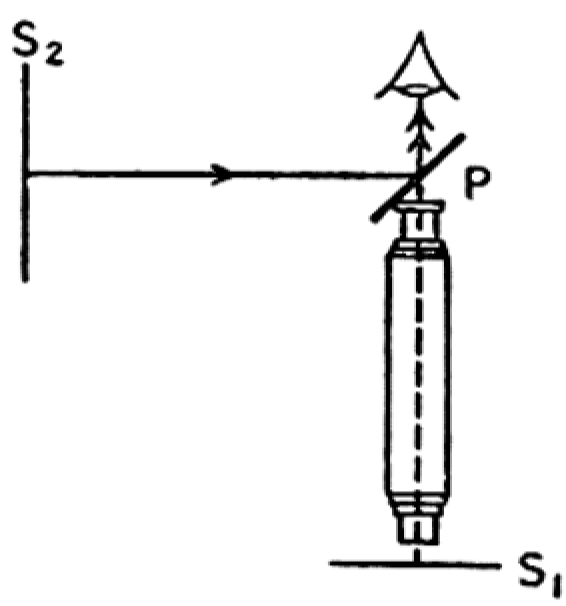


FIG. 664

It is instructive not only to perform the above experiment with an

actual microscope, but also to set up and estimate the magnification of a microscope system consisting of two ordinary short-focus converging lenses—the shorter focus lens being used as the objective. In arranging this model it is useful first to set up the objective alone in front of, say, a millimetre scale and to adjust its position relative to the scale until a highly magnified real inverted image is seen on looking through the objective from a considerable distance. This image, which will be seen to suffer from considerable distortion and chromatic aberration, can be located by a pointer by the elimination of parallax. By doing this, the proper position of the eyepiece can easily be found, because it is only necessary to adjust the position of the second lens to give a magnified virtual image of the pointer. The final image of the scale should then be visible without much further adjustment.

Microscope Objectives.—We have seen that for large magnification without undue lengthening of the microscope tube the objective must have a short focal length. It is easily seen by experiment, as mentioned in the previous paragraph, that the production of a highly magnified real image by a single uncorrected spherical lens is accompanied by aberration and distortions of all kinds. The design of microscope objectives, therefore, is very much concerned with the reduction of these aberrations. A second consideration is the brightness of the final image formed by the microscope. It is evident that for a given distance between the object and the objective, more light from any given portion of the object will enter the objective if the latter has a wide aperture. Notice must therefore be taken of this question in designing objectives even though a wide aperture enhances the aberration problems.

Thirdly, there is the question of the **resolving power** of the microscope. A discussion of the principles of the microscope in terms of geometrical optics suggests that almost unlimited magnification could be produced provided that the problems of aberration and brightness can be adequately solved. The wave nature of light, however, sets another problem which is concerned with the phenomenon of diffraction, the nature of which is discussed in Chapter LII. A spherical light wave emanates from any given point O on the object (Fig. 665), but only a portion of this enters the microscope objective, which is shown in the figure as a hole AB in a screen. As it passes through the aperture, every point on the wave front acts as a source of secondary waves which spread out in all directions on the far side of the aperture, thus causing light to extend sideways into the region outside that enclosed by the dotted lines OA and OB. If light travelled in straight lines, as we suppose it does

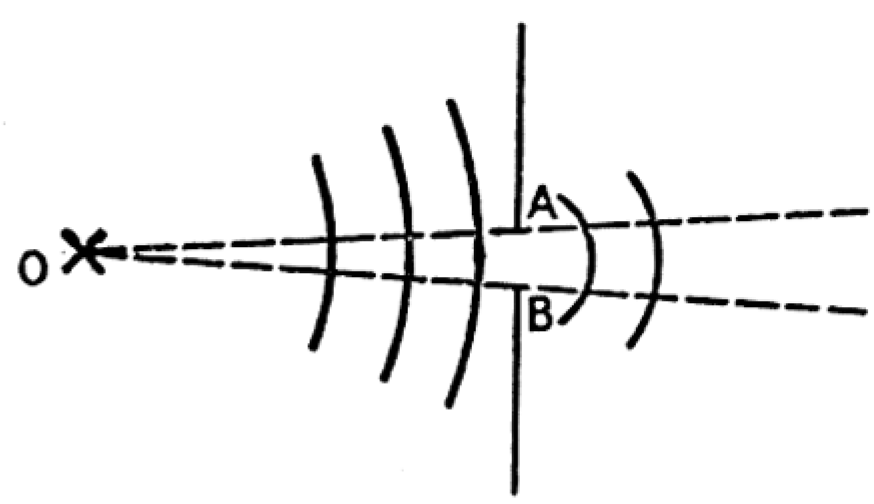


Fig. 665

when we use the conception of rays in geometrical optics, the light from O would be entirely confined between the two straight lines, and a lens placed in the aperture AB would form a perfect point image. The sideways spreading causes the image to cover a finite area instead of being a geometrical point. Actually the interference between light from the various zones of the wave-front in the aperture AB causes an uneven distribution of light in the image of a point object, the effect being a central disc surrounded by concentric alternating dark and light rings. Thus the microscope, and indeed any other optical system, causes each point on the object to be represented by a light pattern of finite size on the image. The effect is not noticeable at moderate magnification, but as the magnification due to an instrument is increased with the intention of seeing more of the detailed structure of the object, a stage is reached at which the patterns due to neighbouring points on the object overlap each other and the separate points are not distinguishable. A further increase in the magnification increases the size of the individual patterns as well as their separation, so that they still overlap and nothing is gained. The only way of achieving such high magnification without loss of resolving power is to take steps to reduce the size of the diffraction pattern produced by each single point on the object so that the overlapping is reduced in consequence. This may be done by increasing the aperture of the objective, because the larger aperture causes comparatively less sideways spreading of the wave-front passing through it. Alternatively, light of shorter wave-length may be used, in which case a given aperture is effectively increased in relation to the wave-length. This has been done by using ultra-violet radiation, which is shorter than visible light, but necessitates the recording of the image by photography. Much more important is the use of a beam of electrons. These particles of negative electricity are corpuscular in many respects, but they do in fact have wave characteristics and therefore produce diffraction, the wave-length being one-thousandth or less that of visible light. In the **electron microscope** the lenses are replaced by coils carrying electric currents which produce magnetic fields and deviate the electron beam in a way comparable with the action of glass lenses on light.

Bearing in mind the questions of magnification, aberrations, brightness of the image and resolving power, it will be realized that the design and construction of microscope objectives is a complicated matter in which it is often necessary to compromise between the factors governing each of the effects. What has been said in Chapter XLVI on the matter of aberrations will lead us to expect a microscope objective to consist of more than one lens so that the deviation is shared between several refracting surfaces. Objectives of moderate power often consist of a pair of achromatic doublets for this reason. With increasing power, the number of components in the system is increased. Use is also made of the principle of aplanatic points (page 783) for producing part of the

large deviation required by a high-power objective. This eliminates spherical aberration and coma.

An important type of objective having high magnifying power and high resolving power is the **oil-immersion achromat.** The first component of such a system (labelled (1) in Fig. 666) has one plane face and one highly curved spherical face. Between the object O and the plane face is a layer of liquid—frequently cedar oil—which has about the same refractive index as the glass of the lens. Thus there is effectively a medium of uniform refractive index between O and the spherical surface of lens (1), and refraction does not occur until this surface is reached. If O is one of the aplanatic points of the spherical surface (page 783) an image of O will be formed at the other aplanatic point I_1, which will be free from spherical aberration in spite of the large angle which the extreme rays make with the principal axis. The lower concave surface of lens (2) is made to have its centre of curvature at I_1, so that no refraction occurs at this face and the upper convex face is given such a curvature as to make I_1 one of the aplanatic points. The image I_2 formed at this point is therefore free from spherical aberration and coma, and a large amount of deviation has been achieved without the accompaniment of these defects. The remaining deviation required for producing the real image is brought about by two doublets (3) and (4), the design of which is made to neutralize the chromatic aberration introduced inevitably by (1) and (2).

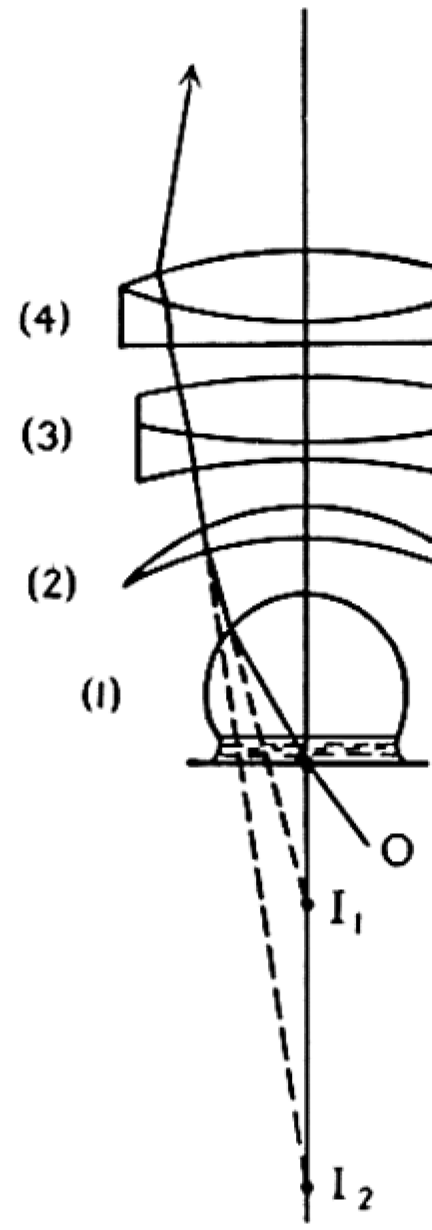


FIG. 666

The device of oil immersion increases resolving power because the wave-length of the light is shorter in the liquid than it would have been in air. It also increases brightness by reducing the loss by reflection at the lower plane face of lens (1) and also by increasing the angle of the cone of light which enters the objective from the illuminating source. This is illustrated in Fig. 667. The light entering the objective aperture AB from the object O is contained in the cone bounded by the rays OA and OB. If, as is frequently the case, the object is viewed by transmitted light sent to it through a converging condensing lens situated below it, then the light passing through the object and eventually entering the objective is contained in the cone bounded by CO and DO, which has a wider angle than the cone AOB because of the refraction at O. Thus, in the absence of

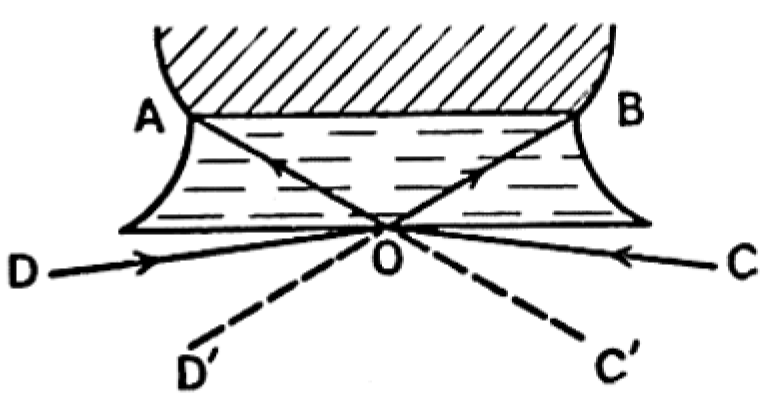


FIG. 667

this refraction, a smaller amount of light contained in the narrower cone between C′O and D′O would have entered AB.

Eyepieces.—An eyepiece (or **ocular**) usually consists of two separate lenses, the one nearer the objective being called the **field lens** and the one nearer the eye the **eye lens.** The purpose of the field lens is to increase the field of view, by which we mean the area of the object which can be seen when the eye is placed in the optimum position. This is illustrated in Fig. 668. Three rays are shown by which the image of the point O′ on the object is formed. Two of these pass through the extreme edge of the objective, while the third passes through its centre. The three

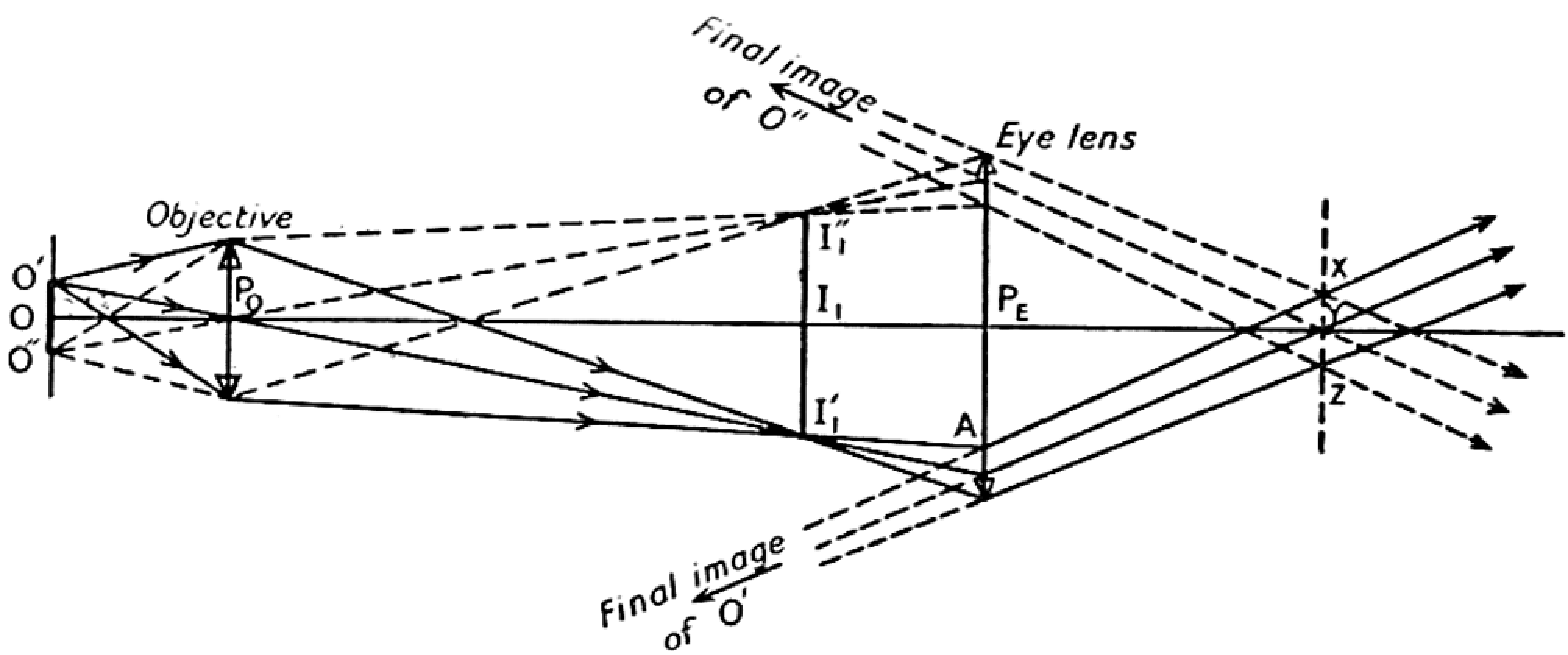


FIG. 668

rays cross at I_1', the real image of O′, and then pass on to the eye lens (in the absence of a field lens), where they are deviated so as to form the magnified virtual image as already explained. The position of the point O′ for which the pencil of rays has been drawn is such that the complete pencil of rays which enters the objective from O′ just succeeds in passing through the eye lens. The same applies to the pencil from the corresponding point O″ on the other side of O, the axial point on the object. For a point on the object just above O′ or below O″ part of the pencil will fall outside the margin of the eye lens and so will not be available for reception by the eye. Such points will, of course, be visible, but their brightness will be less than that of points between O′ and O″. This brightness will evidently diminish as the point moves further from the axis until it becomes invisible when none of the pencil is able to enter the eye lens. This will occur when the point on the object is at a sufficient distance from O to cause the ray marked $I_1'A$ in the drawing to fall outside the margin of the eye lens. Thus the field of view consists of a central region of uniform brightness bounded by such points as O′ and O″, outside which the brightness falls off until it reaches zero at the edge of the field and no point on the object outside this boundary is visible through

the microscope. In actual instruments the field of view is often limited to the central region of uniform brightness by placing a diaphragm in the plane of the first real image $I_1'I_1''$, the diameter of the hole being $I_1'I_1''$. In this way the pencils from points outside OO″ which would be only partially received by the eye lens are entirely cut off, and the edge of the field, being an image of the edge of the stop, is sharp and clearly focused. The rays from O′ are shown continued past the eye lens so as to cross the common axis of the lenses, and the corresponding rays are shown as dotted lines from the point O″. At the plane XYZ the beam issuing from the eye lens has its smallest cross-section, and the area of this, represented by XZ on the drawing, is called the **exit pupil.** A stop is often fixed here and is called the **eye ring,** since it indicates the best position for placing the eye in order to see the whole field of view presented by the instrument. If the eye were placed elsewhere along the axis the limited size of its pupil would prevent the more oblique rays from entering it, with the result that the outer part of the field would be cut off. The position of the exit pupil can be found experimentally for any particular instrument by illuminating the front of the objective and moving a white card towards or away from the eye lens until the patch of light on it has its minimum diameter. The patch is then the exit pupil XZ. It is the real inverted image of the objective formed by the eye lens, because the two rays from the top of the objective intersect at Z and the pair from the bottom intersect at X. The point Y is the image of the centre of the objective P_O.

Returning to the question of increasing the field of view by a field lens, we can illustrate the principle by Fig. 669. In this, a pencil of rays is shown converging to a point I on the first real image. In the absence of the field lens the pencil would follow the dotted lines. Suppose that the eye lens is not sufficiently wide to receive this pencil, with the result that the point on the object from which the rays originated would be outside the field of view. The field lens is placed in the plane of the image and the pencil is bent as shown to pass through the eye lens, thus extending the field of view. A field lens placed in this position has no effect on magnification because it does not move the position of I, since this point is in the plane of the lens ($u=0$ and $v=0$ in the lens equation). In practice, this type of eyepiece, which often has a chromatic doublet as the eye lens, has the disadvantage that specks of dust on the surface of the field lens are seen clearly in focus because the eye lens is focused on this surface.

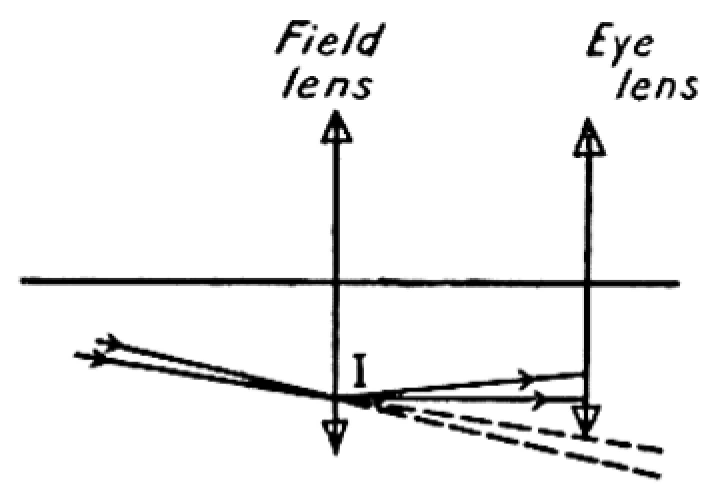


FIG. 669

A type of eyepiece known as the **Huygens eyepiece** is often used in microscopes. It is shown in Fig. 670 as consisting of two plano-convex lenses of the same glass. If the focal length of the eye lens is f, that of

the field lens is made to be $3f$, and the distance between the lenses is $2f$. This fulfils the condition for achromatism of two separated lenses referred to on page 831, namely that the separation of the lenses shall be equal to their mean focal length. The field lens intercepts the rays from the objective before they form the first real image I_1I_1', which therefore acts as a virtual object for the field lens and gives rise to the real image I_2I_2'. This is arranged to be in the first focal plane of the eye lens, so that the rays emerge from the latter parallel to each other and the final image is at infinity. Spherical aberration is kept small by the equal sharing of the deviation between the four lens surfaces.

The **Ramsden eyepiece** (Fig. 671) also consists of two plano-convex lenses of the same glass, but their focal lengths are made equal. According

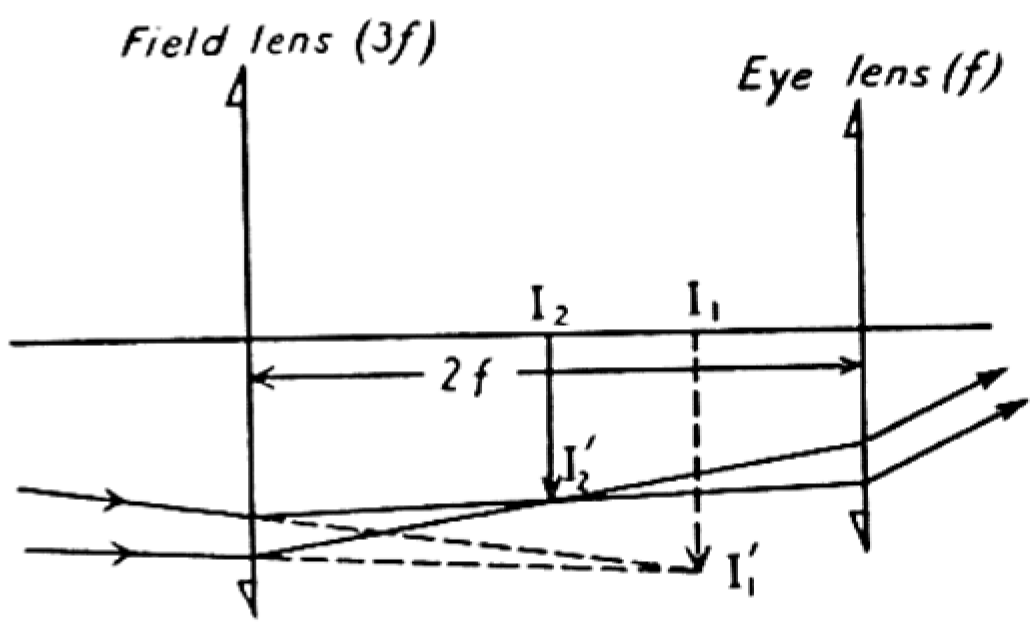


FIG. 670

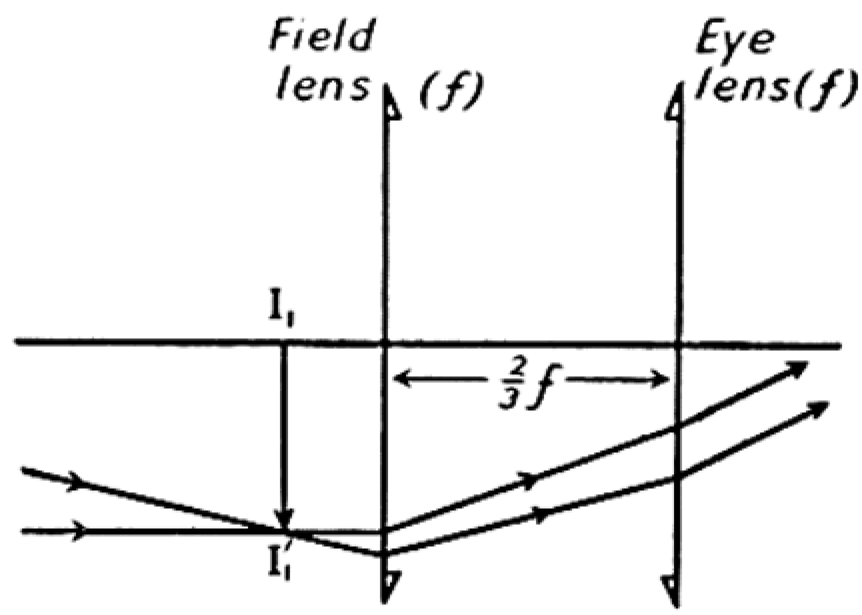


FIG. 671

to the rule for achromatism the separation of the lenses should therefore be equal to the focal length of either of them. However, if the final image is to be at infinity, this would mean that the first real image would be in the first focal plane of the eye lens and would therefore coincide with the surface of the field lens and, as already mentioned, this causes dust and scratches on this lens to be clearly visible. As a compromise, therefore, the separation is reduced to two-thirds of the focal length of either lens, and the first real image I_1I_1' is arranged to be in front of the field lens as shown in Fig. 671, but within its focal distance. One advantage of this type of eyepiece lies in the fact that crosswires or an eyepiece scale can be placed in the microscope tube at the plane I_1I_1', and the image of these is seen superimposed on that of the object which is being viewed. The Ramsden eyepiece is, therefore, frequently used in travelling microscopes and cathetometers.

Illumination of the Object.—In serious microscopy the question of the illumination of the object is as important as the design of the objective and eyepiece. Owing to the very high magnification it is necessary to concentrate as much light as possible on to the object, and this is done by placing under the stage and coaxial with the microscope a high-power converging lens system known as the **condenser** through which light is passed from below. An inclined mirror below the condenser is often

used to direct the light through it from a lamp placed on the table near the instrument. A microscope condenser may be regarded as an objective used in reverse. It produces a small, and therefore bright, image of the source in the plane of the microscope slide as indicated in Fig. 672, and light from this image enters the objective. It is found that the avoidance of aberrations in the condenser is almost as important as in the objective.

When very small particles are being viewed at high magnification, *e.g.* when Brownian motion is being examined, it is often more satisfactory to use **dark-ground illumination.** In this, the condenser concentrates light on the object, but the light itself does not subsequently pass directly into the microscope. There are various ways of designing a condenser

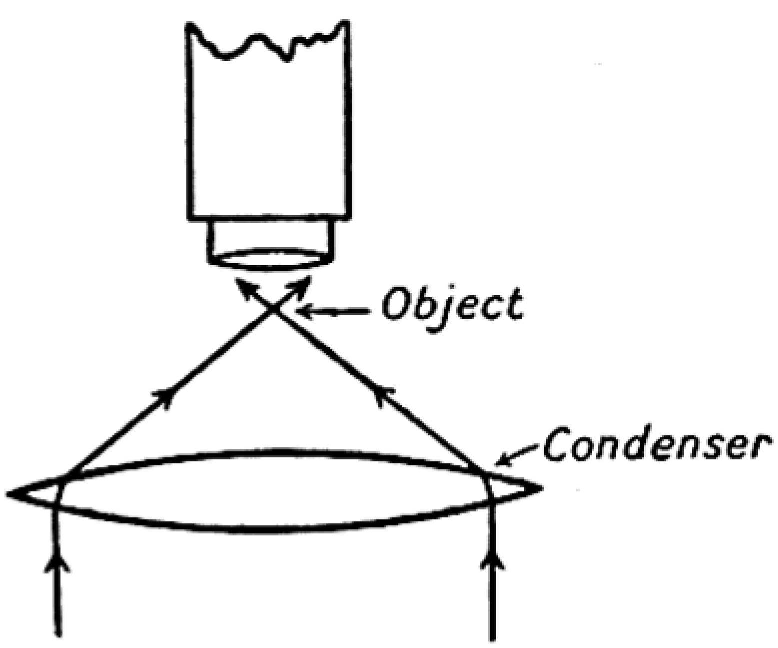


FIG. 672

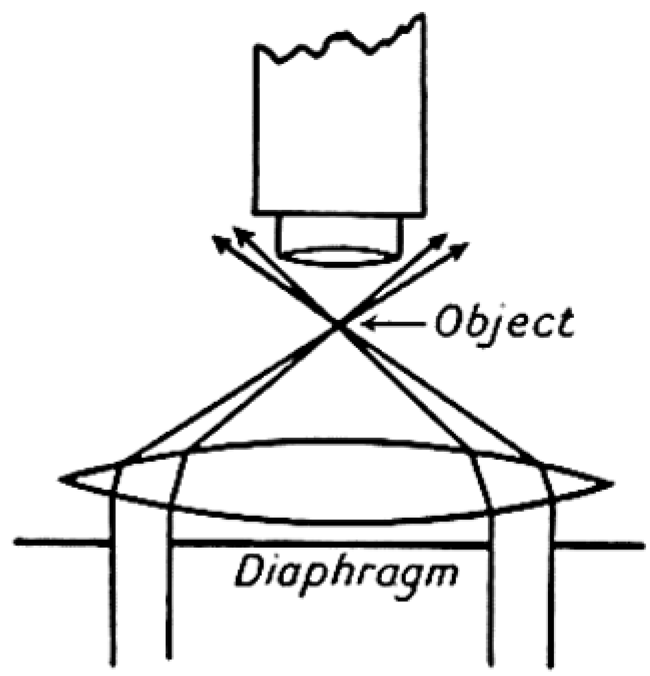


FIG. 673

to achieve this, one of the simplest principles being to block out the central rays passing through a highly converging lens system by means of a diaphragm so that a hollow cone of light is produced as in Fig. 673. If no object is present, the field of view in the microscope is dark, but a small object placed at the apex of the cone will *scatter* light into the microscope and appear light against the dark background.

3. THE TELESCOPE

The Astronomical Telescope.—In general it may be said that microscopes are used to produce magnification of objects which are too small to be seen in detail even when held at the near point of the naked eye, whereas telescopes are used to produce magnification of objects which appear small to the naked eye by virtue of their distance from it. In a microscope there are, as we have seen, two stages of magnification—the first by the objective and the second by the eyepiece. In the telescope, just as in the microscope, the objective produces a real image of the distant object, but this is necessarily smaller than the object itself because it must be situated within the telescope tube, and its distance from the objective is therefore many times less than the distance of the object from this lens. The eyepiece acts as a simple microscope to produce magnification

of the real image. The final image, especially if it is formed at the near point of the eye, may be very much smaller than the distant object, but it must be remembered that it is the angle which it subtends at the eye which determines its apparent size. The objective as well as the eyepiece contributes to the process of angular magnification by effectively bringing the object nearer to the eye. Indeed a single converging lens can have the effect of a telescope without the addition of an eyepiece. This is illustrated in Fig. 674. Three parallel rays are shown falling on the lens from the top O′ of a very distant object, the bottom of which (O) is situated on the principal axis of the lens. Thus the real, diminished, inverted image II′ is formed in the focal plane of the lens. An eye is

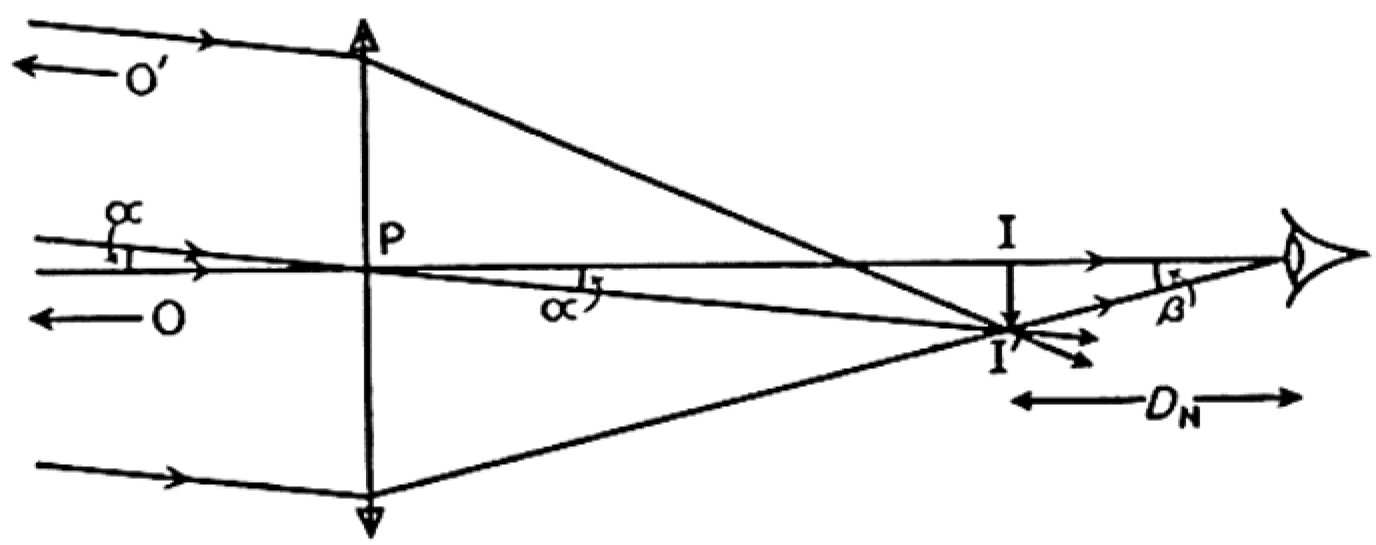


FIG. 674

placed beyond this image and, in order that the image shall appear as large as possible while still being sharply in focus, it should be situated at the near point of the eye, *i.e.* a distance D_N from the eye. The angle β subtended at the eye by the image is then equal to $\frac{II'}{D_N}$. To find the angle which OO′ would subtend at the eye if viewed direct, we can suppose that the eye is placed at P. This shift of its position, being small compared with the distance of OO′, will have a negligible effect on the angle, which is marked α in the drawing. Evidently α is equal to $\frac{II'}{PI}$, where PI is numerically equal to the focal length of the lens, so that the angular magnification is given by

$$\frac{\beta}{\alpha}=\frac{II'}{D_N}\div\frac{II'}{PI}$$

$$=\frac{PI}{D_N}$$

An angular magnification of greater than unity will be achieved, therefore, if the focal length is greater than the value of D_N for the observer's eye. Thus with an eye having the standard value of 25 cm. for D_N and a lens of power +1 dioptre, for which PI will be equal to 100 cm., a magnification of 4 is obtained.

Suppose that a second converging lens, of focal length f_E, is used as a magnifying glass to view the real image II′. This will increase the angle subtended at the eye by a factor equal to the angular magnification which it produces, which, if the final image is at infinity, will be $\frac{D_N}{f_E}$. Thus the final angular magnification is greater by this factor than that obtained when the first lens alone is used, and it is therefore equal to $\frac{PI}{f_E}$ or $\frac{f_O}{f_E}$ if we denote the focal length of the first lens (the objective) by f_O. Thus by considering the action of the objective and eyepiece separately we have arrived at a formula for the total angular magnification, or magnifying

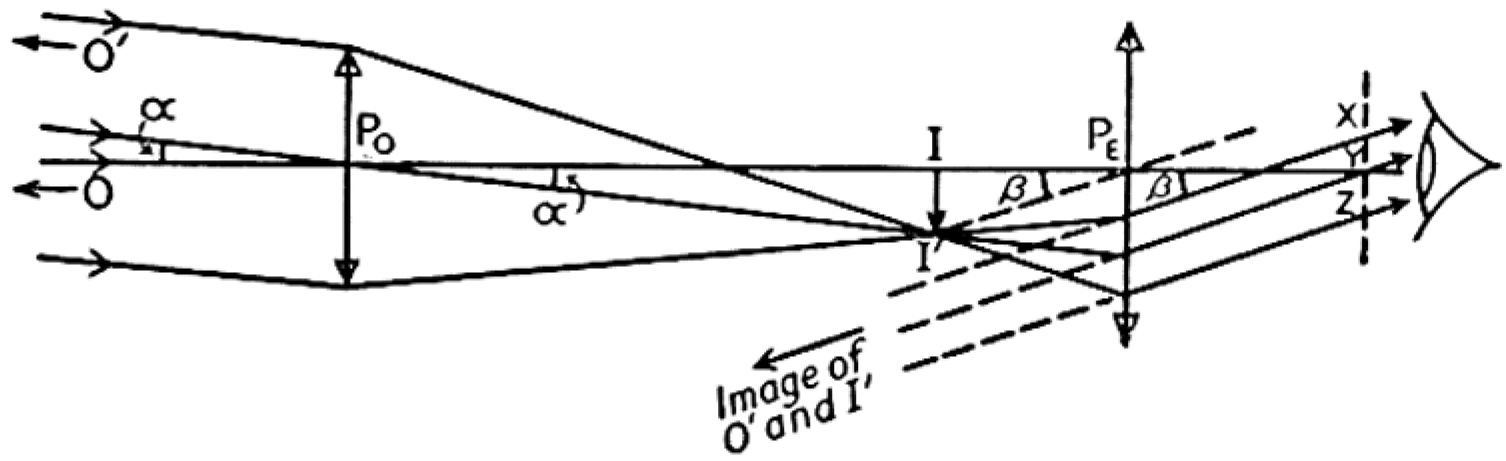


FIG. 675

power, of a telescope when it is in **normal adjustment,** *i.e.* when the object and final image are both at infinity. We shall see, however, that the formula is easily obtained directly from the ray diagram which is shown in Fig. 675.

In drawing the ray diagram of an astronomical telescope in normal adjustment we take three parallel rays which are supposed to come from the top (O′) of the very distant object. These are converged to a point I′ (the centre ray being undeviated as it passes through the centre of the objective P_O). If O, the bottom of the object, is on the principal axis, the first real image of the object is II′, which is perpendicular to the axis. The distance IP_O is the focal length f_O of the objective and, since the final image is to be at infinity, IP_E must be equal to f_E, the focal length of the eyepiece, the point I being the principal focus of both lenses. The rays from the objective cross at I′ and reach the eyepiece. In order to find their direction after refraction by this lens it is only necessary to draw the broken construction line $I'P_E$, which represents a ray from I′ striking the lens at its centre P_E and passing through undeviated. Produced backwards, it would pass through the final image of I′, which is at infinity. Therefore all other rays from I′ will be deviated by the lens so as to travel away from it parallel to the direction $I'P_E$ because they also, if produced backwards, would pass through the image of I′.

Remembering that the base of the object (and therefore of the final image) is on the axis of the telescope, we see that the angle subtended at

the eye by the final image is the angle β between the axis and the emergent rays which originally came from O′. In particular the angle $\widehat{IP_EI'}$ is equal to β and can be expressed as $\frac{II'}{IP_E}$. Furthermore, the angle α which OO′ would subtend at the eye if it were viewed direct is equal to $\widehat{IP_OI'}$, which is equal to $\frac{II'}{IP_O}$. Consequently the angular magnification is given by

$$\frac{\beta}{\alpha}=\frac{IP_O}{IP_E}$$

For normal adjustment, as we have already seen, $IP_O=f_O$ and $IP_E=f_E$, so that

$$\frac{\beta}{\alpha}=\frac{f_O}{f_E}$$

which is the equation arrived at previously by a consideration of the separate angular magnifications produced by the two lenses.

The distances between the lenses is evidently equal to (f_O+f_E) for normal adjustment.

If the object is not infinitely distant, the three rays from O′ will be slightly divergent and the image II′ will move slightly to the right beyond the focal plane of the objective. If, in addition, the eyepiece is placed in such a position that the final image is nearer than for normal adjustment, then IP_E must be less than f_E. The student will find it instructive to draw a ray diagram of this case, finding the position of the final image by using the usual geometrical construction at the eyepiece and drawing the paths of continuous rays right through the instrument as in Fig. 675.

Example.—*An astronomical telescope consisting of two thin converging lenses gives an angular magnification of* 5 *when in normal adjustment and the lenses are* 48 *cm. apart. What are the focal lengths of the lenses? What would be the distance between the lenses if the object is brought to a distance of* 2 *metres from the objective and the final image is at infinity?*

For normal adjustment the object and final image are both at infinity, the distance between the lenses is equal to the sum of their focal lengths (f_O+f_E), and the angular magnification is equal to f_O/f_E. Since both the focal lengths are positive we can write

$$f_O+f_E=48$$

and

$$\frac{f_O}{f_E}=5$$

Combining the two equations we obtain

$$f_O=40 \text{ cm.}$$

and

$$f_E=8 \text{ cm.}$$

When the object is 200 cm. from the objective it forms the first real image at a position v given by

$$\frac{1}{v}+\frac{1}{200}=\frac{1}{40}$$

which gives

$$v=50 \text{ cm.}$$

In order that the final image shall be at infinity the first real image must be at the principal focus of the eyepiece, *i.e.* 8 cm. from it. Therefore the distance between the lenses is now (50 + 8) or 58 cm. That is to say, the telescope has been lengthened by 10 cm.

We shall now discuss the case of normal adjustment (Fig. 675) a little more. The ray from O′ through P_O crosses the axis of the telescope at Y, which can easily be seen to be the image formed by the eyepiece of the centre of the objective P_O. The plane XYZ is perpendicular to the axis. If we drew a set of parallel rays similar to those from O′ from any other point on the object (*e.g.* from O or from a point as far below the axis as O′ is above it), we should find that all rays passing through the top edge of the objective finally pass through Z, and all rays from the bottom pass through X. In fact, of course, XYZ is the image of the objective, and is called the exit pupil as it is in the compound microscope (Fig. 668, page 874). An eye placed at the exit pupil can see the whole field without changing its position, and the eye ring is placed here.

It is instructive to draw the path of two rays passing through the edge of the objective from an axial point on the object as in Fig. 676. These are initially parallel to the axis. They cross over at I, the common principal focus of the lenses, and are then made parallel by the eyepiece. The position of the exit pupil (the image of the objective AB) is easily found by means of the broken lines AP_E and BP_E from the edge of the objective passing through the centre of the eyepiece. From a consideration of similar triangles it is evident that

$$\frac{AB}{XZ}=\frac{AB}{CD}=\frac{IP_O}{IP_E}=\text{angular magnification}$$

Thus, when the telescope is in normal adjustment, the ratio of the diameter of the objective to the diameter of exit pupil (which can be found by locating it on a white screen when the objective is illuminated) is equal to the angular magnification. This fact can also be proved by using the lens equation for finding the position and size of the image of the objective formed by the eyepiece.

Suppose that the width of the exit pupil (*i.e.* of the emergent parallel beam in Fig. 676) is less than the width of the pupil of the observer's eye. Let the diameter of the objective AB be now increased, keeping all other factors constant. This will cause more of the light emitted by a given small area of the object to be collected by the objective, so that more light emerges from the instrument into the eye. This is made evident

by the increased width of the emergent beam. However, as AB is increased, a stage will be reached at which the emergent beam completely fills the eye pupil. Further increase of AB, although it causes still more light to enter the instrument, does not increase the brightness of the image because the additional light now falls outside the eye pupil. In other words, the additional area of the objective is ineffective as regards further increase of brightness. When the exit pupil is the same size as the pupil of the eye, so that the maximum brightness is being obtained, it is clear that the ratio of the diameter of the objective to that of the eye pupil is equal to the angular magnification (m) and the ratio of their areas is m^2. On account of this ratio of areas, the amount of light entering the telescope from the object is m^2 times the amount of light which would enter the eye pupil if the object were viewed directly from the same distance. But the area of the image received on the retina when the telescope is used is also

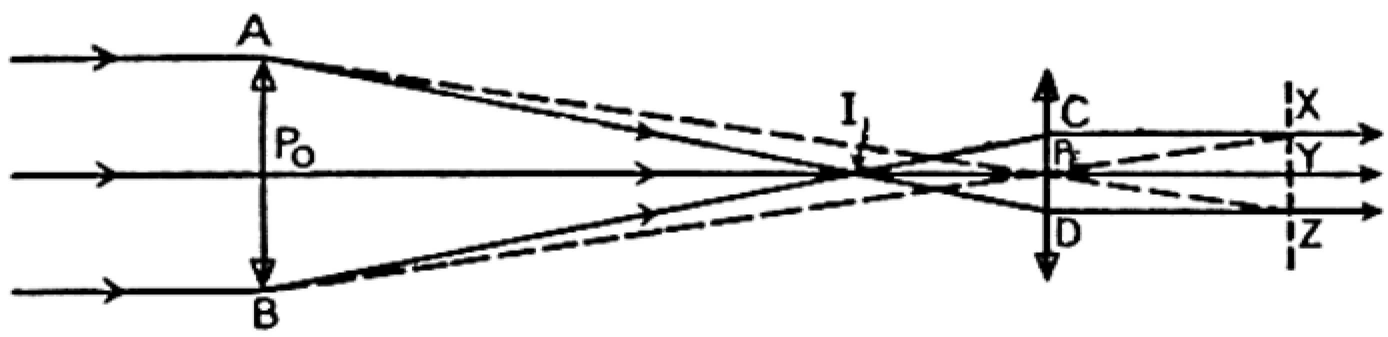


FIG. 676

m^2 times as large as when the object is viewed with the naked eye, so that the greater amount of light is distributed over an area which is greater by the same factor, with the result that the apparent brightness of the object is the same whether the telescope is used or not. In practice, of course, reflection at the lens surfaces and absorption in the glass will make the brightness less when the telescope is used. The foregoing argument refers to an *extended object*, *i.e.* one which subtends a finite angle at the eye.

When a very distant object such as a star is viewed it is effectively a point source, and the retina of the eye receives a small diffraction pattern both when the object is viewed directly and when a telescope is used. In the latter case the amount of light received is greater in the ratio of the area of the objective to that of the eye pupil, but the size of the patch of light on the retina is not changed in the same ratio. Thus in such a case the telescope increases brightness rather than apparent size, and telescopes with wide objectives are, of course, used by astronomers to view and to photograph stars which are so distant as to be invisible to the naked eye. The telescope, so to speak, artificially increases the light-gathering power of the eye. Suppose that a certain star is just visible in a particular telescope, and that another instrument is then used with an objective of twice the width of the first. This instrument will collect four times as much light from the given star, which would therefore be still just visible if its distance were doubled (on account of the inverse square law, page 901). Thus the range of a telescope is proportional to the diameter of its objective.

The brightness of the background is not increased in the same way as that of the star itself when a telescope is used because it acts as an extended object, the brightness of which is, as explained in the previous paragraph, not increased by the telescope. Thus contrast is increased by the use of the telescope.

The **resolving power** (page 953) of a telescope is increased by increasing the diameter of the objective since, by doing so, the size of the diffraction pattern due to a point object is reduced.

The question of the **field of view** of the telescope can be discussed in the same way as for a microscope (page 874). Remembering that in Fig. 675 parallel incident rays all come from the same point on the object, we can say that a point on the object will be visible in the telescope if the rays from it eventually pass through the eyepiece. Suppose that, instead of (as in Fig. 675) the bottom of the object being on the axis of the telescope, the object extends indefinitely in all directions from the axis outwards. There will then be a region of the object, with the axial point as centre, which is visible with uniform brightness, because all the rays which enter the objective from any point in this region also pass through the eyepiece. Immediately outside this region only a part of the pencil reaching the plane of the eyepiece actually passes through this lens, the remaining part being outside the margin of the lens. This can be appreciated by imagining the angle α in Fig. 675 to be gradually increased. In this peripheral region of the field of view, therefore, the brightness falls off with increasing distance from the axis until it becomes zero when the whole of the pencil falls outside the area of the eyepiece. This is the outer limit of the field of view.

As with the microscope, the field of view can be limited to the central area of uniform brightness by placing at the first real image a stop of such a diameter that it prevents all but the complete pencils from entering

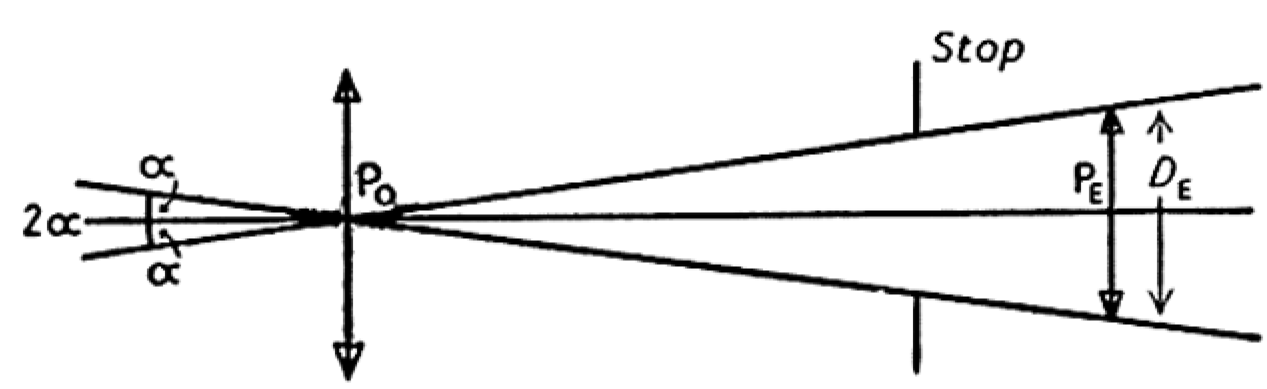


FIG. 677

the eyepiece. Alternatively the stop can be made wider so as to pass those partial pencils whose central ray strikes the eyepiece. The field of view under these conditions can be calculated quite simply. It is expressed in terms of the angle 2α subtended by the visible area of the object, and Fig. 677, in which the central rays of two extreme pencils are drawn, shows that

$$2\alpha = \frac{D_E}{P_O P_E}$$

where D_E is the diameter of the eyepiece. For normal adjustment $P_OP_E = (f_O + f_E)$, and the field of view is given by $\frac{D_E}{f_O + f_E}$. This is called the field of view "in the object space." The field of view in the image space is the angle subtended at the eye by the image of the visible area, and will be equal to 2α multiplied by the angular magnification. It will be realized that, on account of the normally small value of D_E and the large value of $(f_O + f_E)$, the field of view of a telescope is very much smaller than that of, say, a box camera, in which the corresponding distances are, respectively, the transverse dimensions of the film and the distance from the lens to the film.

In practice a telescope is always fitted with a compound eyepiece of one of the types described in connection with the microscope (pages 874–876). These have the effect of reducing aberrations and of increasing the field of view (page 875). Steps are also taken to minimize aberrations in the objective, which often consists of an achromatic doublet such as that shown in Fig. 678, with the back surface plane or at any rate less curved than the front surface. The deviation is more equally shared between the two faces by this means, thus reducing spherical aberration.

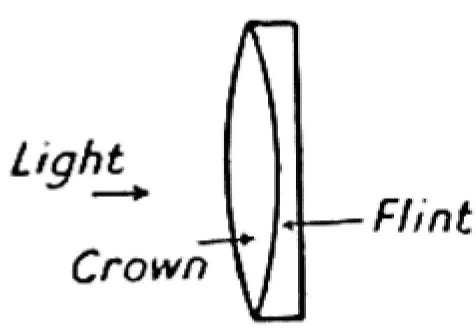


FIG. 678

Determination of Angular Magnification.—A simple and direct method of estimating the angular magnification due to a telescope under any particular conditions consists in arranging matters so that the object which is being viewed is looked at directly with one eye, while the other eye is looking through the instrument. By suitable adjustment the image can be superimposed on, or made to appear alongside, the object, so that the ratio of corresponding linear dimensions on the two can then be estimated. This ratio is equal to the angular magnification.

Another method makes use of the fact (page 881) that for normal adjustment the ratio of the diameter of the objective to that of the exit pupil is equal to the angular magnification. Diffused light is made to enter the objective and a white screen is moved about near the eyepiece until it receives a clear image of the aperture of the objective formed by the eyepiece. This is the exit pupil, sometimes called **Ramsden's circle,** and its diameter can be measured.

The Terrestrial Telescope.—The astronomical telescope, consisting of two lenses or systems of lenses, is evidently unsuitable for the observation of terrestrial objects because of the inversion which it introduces. A telescope in which the inversion is eliminated is known as a terrestrial telescope. An obvious way of doing this is to introduce a converging lens between the first real image and the eyepiece. If the distance of this erecting lens beyond the first real image is twice the focal length of the former, it will form an erect image of the same size and at the same distance beyond the lens. This is illustrated in Fig. 679, in which rays coming

from the objective are shown passing through the first image I_1I_1' to the erecting lens, and so through the second upright real image I_2I_2' to the eyepiece which produces the usual magnified virtual image. In practice

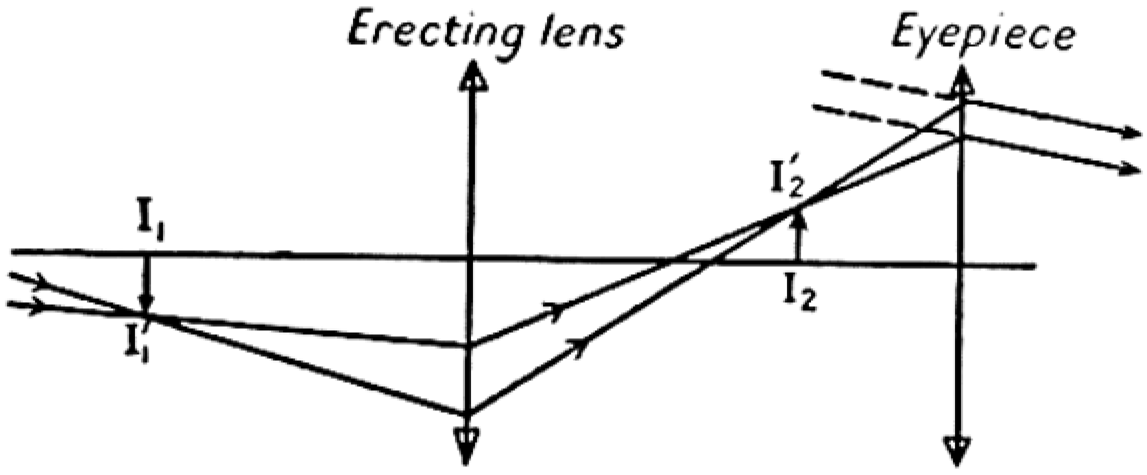


FIG. 679

the erecting system often consists of two separated lenses, thus enabling aberrations to be reduced. A terrestrial telescope using a lens or lenses for inversion obviously suffers from the disadvantage that the length of the instrument must exceed that of the corresponding astronomical telescope by the distance between the first and second real images.

Prism Binoculars.—Prism binoculars consist of a pair of telescopes, each of which is made up of an objective and an eyepiece and is, therefore, effectively an astronomical telescope. In passing from the objective to the eyepiece the light is reflected by two right-angle glass prisms, the action of which is described on page 752. In Fig. 680 the action of the prisms on the path of a ray travelling along the axis of the objective is shown.

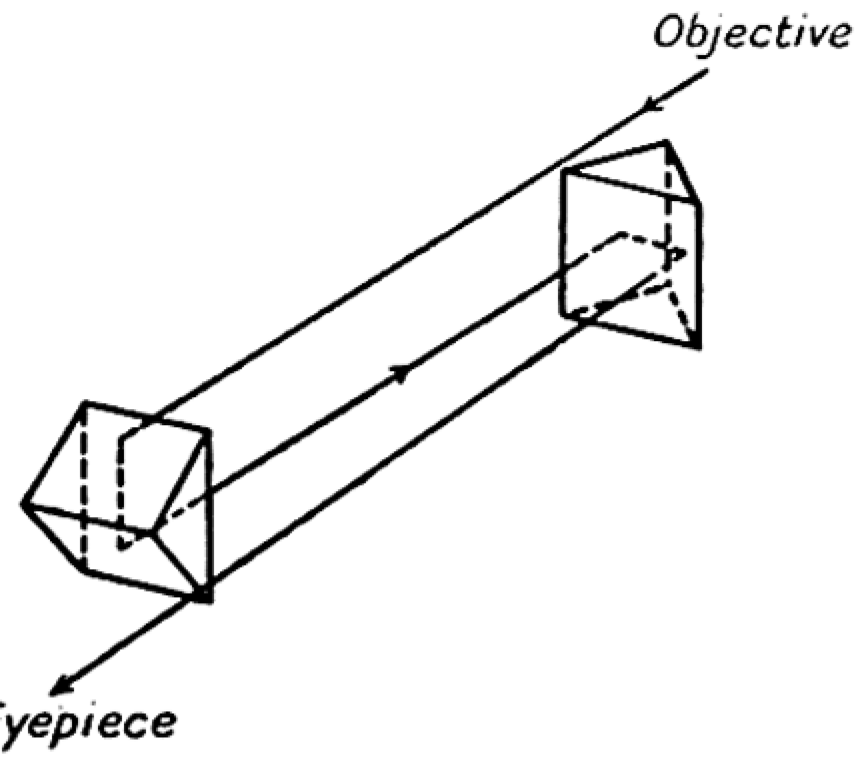


FIG. 680

The arrangement is equivalent to an ordinary telescope in which the distance between the lenses is equal to the length of the path of the light in Fig. 680. It is therefore possible to use an objective with a comparatively long focal length, thereby achieving the magnification and performance of a telescope whose length would make it cumbersome but for the use of the prisms.

Each prism produces an inversion, as can be seen from Fig. 681. The central ray (labelled (1)) of a pencil from a given point on the object would form an image at, say, I_1 in the absence of the prism. The prism deviates the central ray through 180° and the image is formed at I_2. A similar ray (2) from another point forms an image at I_2', which is above I_2, whereas if the prism had not been present it would have been at I_1', below I_1. The image I_2I_2' is therefore inverted relative to I_1I_1'. In Fig. 680 it will be noticed that the 90° edges of the two prisms are at right angles to each other. One of the prisms,

therefore, cancels the inversion produced by the objective in a vertical plane, while the other cancels the lateral inversion. Prism binoculars or "field glasses" are often made to give a magnification of 6 or 8.

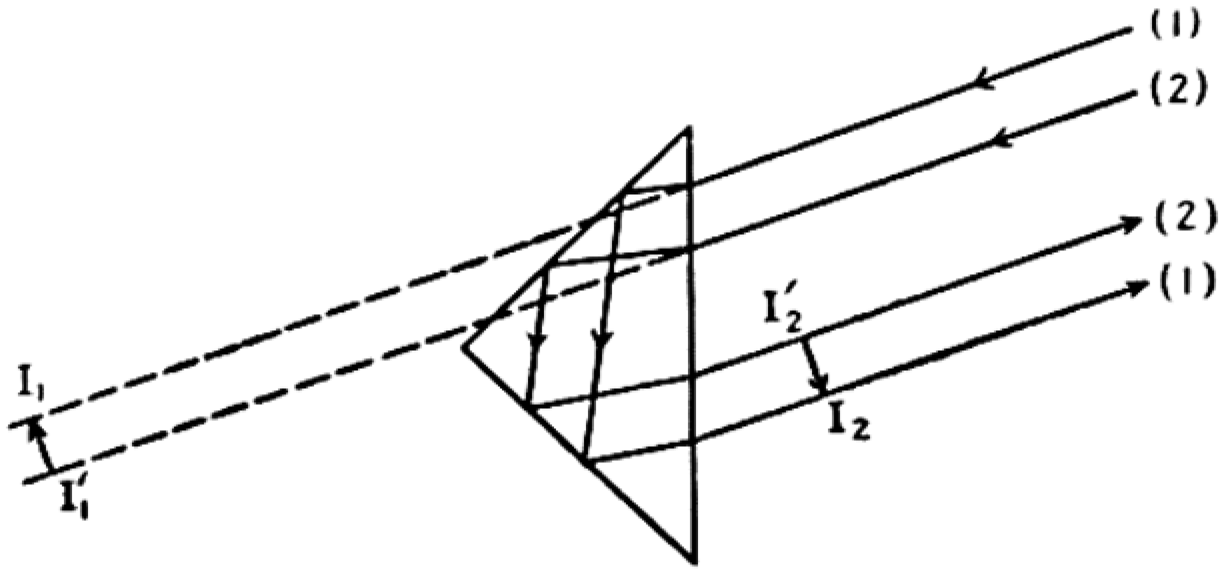


FIG. 681

Galileo's Telescope.—In this instrument the objective is a converging lens and the eyepiece is a diverging lens placed so as to intercept the light from the objective before it forms a real image. The result is an image with angular magnification but without inversion. Fig. 682 is a ray diagram corresponding to Fig. 675 for the astronomical telescope. The

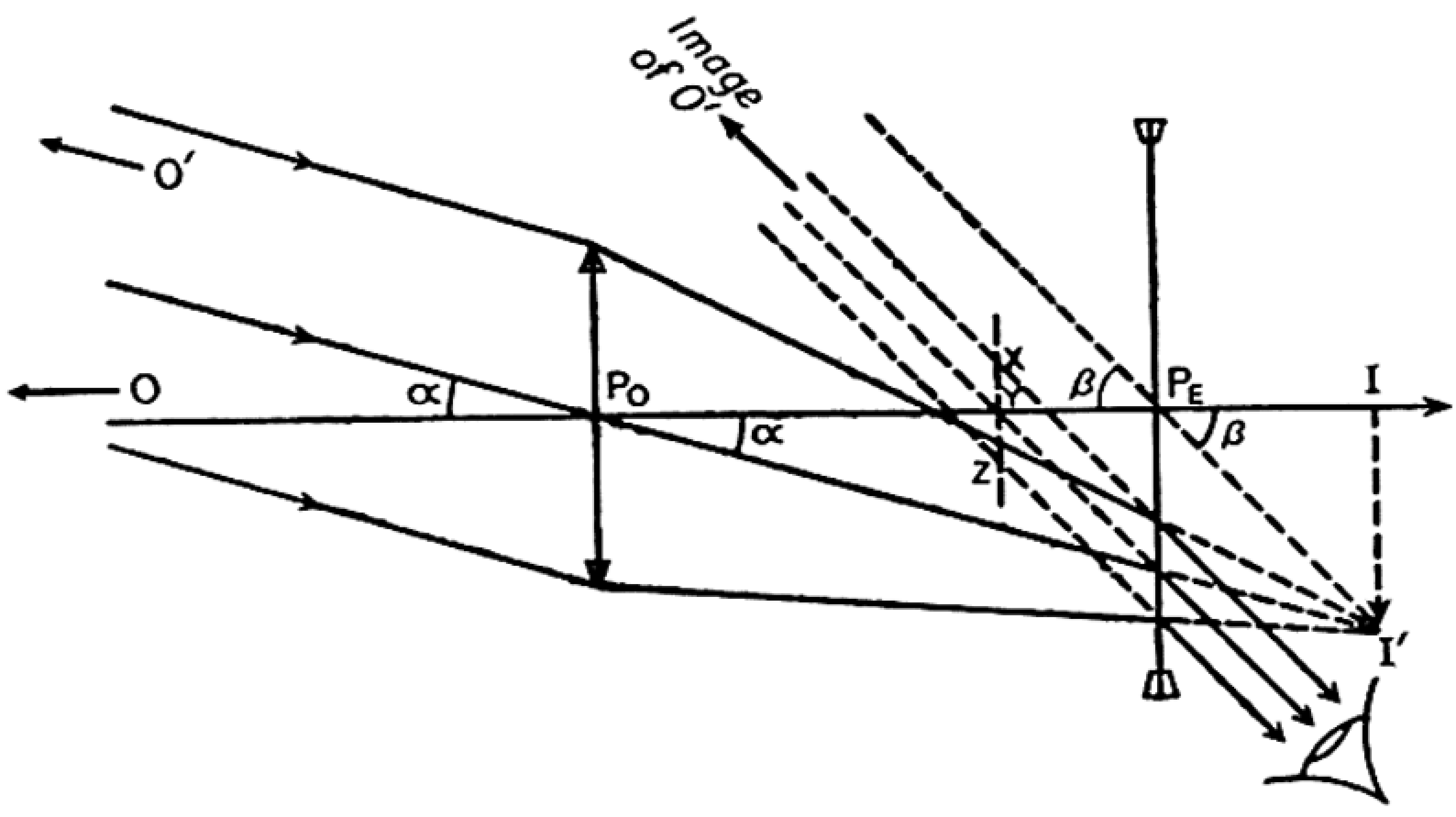


FIG. 682

telescope is supposed to be in normal adjustment. Therefore the object OO' is very distant, and its image II' formed by the objective (focal length f_O) would be in the second focal plane of the objective if the light were not intercepted by the eyepiece. Furthermore, the eyepiece (of focal length f_E) is so placed that II' is in its focal plane, which means that rays converging to any point on II' emerge from the eyepiece parallel to each other. Three such rays coming originally from the top of the object O' and made to converge to I' are shown in the drawing. After passing through the eyepiece their directions must be parallel to the line P_EI'. Thus the final image of O' is at infinity on these rays produced backwards,

and that of O is at infinity on the axis. The angular magnification is equal to β/α, where $\alpha = \frac{II'}{IP_O}$ and $\beta = \frac{II'}{IP_E}$, so that

$$\frac{\beta}{\alpha} = \frac{IP_O}{IP_E}$$

$$= -\frac{f_O}{f_E}$$

for normal adjustment, the negative sign appearing because f_E is regarded as negative on the R.P. sign convention. Evidently the distance between the lenses for normal adjustment is equal to $(IP_O) - (IP_E)$, which is the difference between the numerical values of the focal lengths and can be written $(f_O + f_E)$ since f_E is negative.

The equivalent of the exit pupil in the astronomical telescope (XYZ in Fig. 676) is to be found at XYZ in Fig. 682. All rays leaving the eyepiece must have directions which pass through this area. The pupil of the eye cannot, of course, be placed at XYZ so as to obtain the largest field of view, as is possible with the astronomical telescope. The best that can be done is to place the eye as near as possible to the eyepiece, and then the field of view is evidently determined by the diameter of the eye's pupil. If this is equal to e, then the field of view, estimated in the same way as for the astronomical telescope (page 883), will be represented by the angle 2α (Fig. 683), which is equal to $\frac{e}{f_O + f_E}$. This is smaller than for a comparable astronomical telescope on account of the smallness of e.

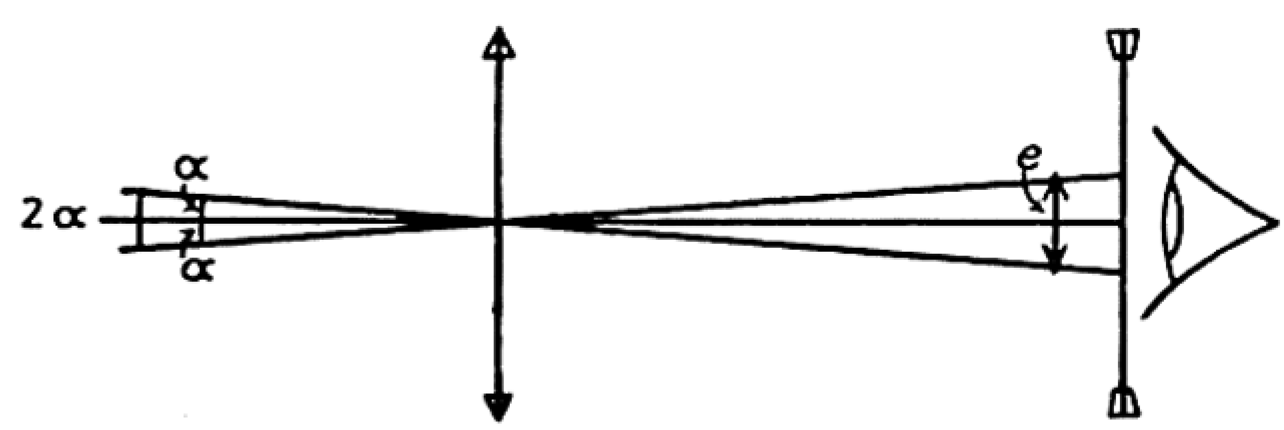


FIG. 683

The main advantages of the Galilean telescope are its comparatively short length and its erect image. One of its commonest forms is to be found in **opera glasses.**

Reflecting Telescopes.—Newton was unaware of the possibility of reducing the chromatic aberration of lenses by using two or more different types of glass in contact (page 827), and he proposed to eliminate this defect from telescopes by using a concave mirror in place of a convex lens for the objective. Reflection is not, of course, subject to chromatic aberration. The principle of Newton's reflecting telescope is shown in

Fig. 684. The long-focus concave mirror converges light from a point on a distant object and would produce a real image at I. For convenience of viewing and magnifying this image by means of the eyepiece, a plane reflector intercepts the converging light and causes the image to be

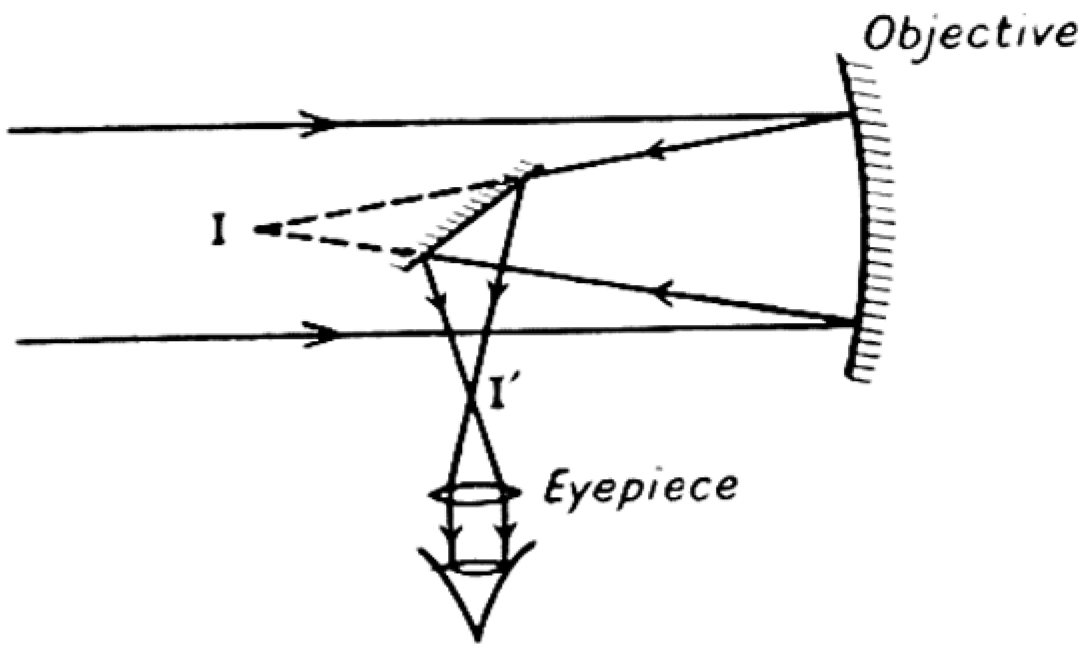


FIG. 684

formed at I′. The eyepiece then acts in exactly the same way as in a refracting telescope.

Another arrangement, known as the Cassegrain reflecting telescope, is shown in Fig. 685. In this, the light proceeding to the first real image I is intercepted by a convex mirror and made to pass through a hole in the centre of the objective and to form the real image at I′.

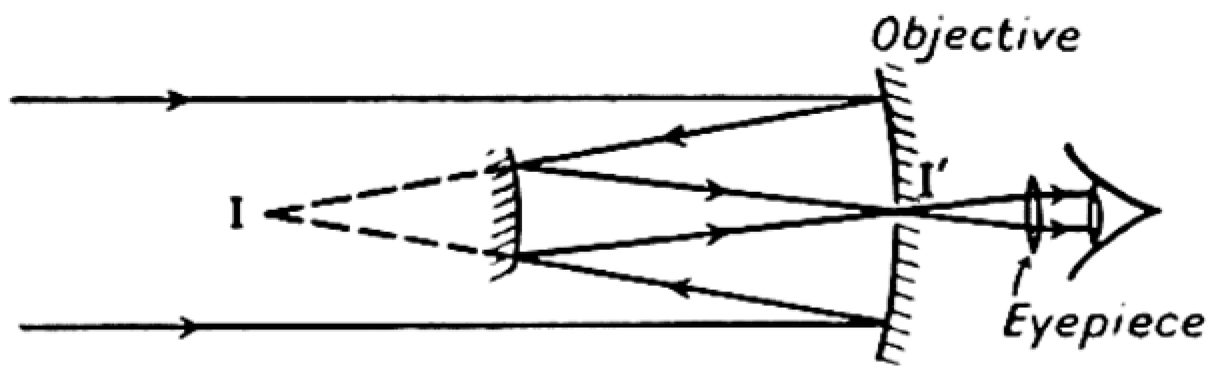


FIG. 685

Since the discovery of the achromatic doublet, most telescopes of ordinary size are refractors. It has been pointed out (page 882), however, that the range of a telescope used in astronomy (*i.e.* the greatest distance at which a telescope can render visible a star of given intensity) increases with the diameter of the objective. The difficulties associated with the construction and use of large refracting objectives are very great. A large mass of glass must necessarily be used and the lens tends to bend under its own weight. Also the necessary uniformity of refractive index throughout the glass is difficult to obtain. These two troubles are not present when a reflector is used because it can be supported all over its back surface, and the front surface is "silvered" (nowadays with aluminium). In addition to the absence of chromatic aberration in a reflecting telescope objective, spherical aberration for points on the object near the axis of the mirror can be made very small by giving the reflecting surface a parabolic

shape (page 735). For the reasons given in this paragraph, therefore, telescopes for photographing the heavens, such as the 100-in. telescope at Mt. Wilson and the 200-in. instrument at Mt. Palomar, are reflecting telescopes.

4. THE CAMERA

A camera consists essentially of a converging lens through which light from the scene or object to be recorded passes into a light-tight enclosure and there forms a real image on a film or plate coated with a light-sensitive emulsion. The arrangement is illustrated schematically in Fig. 686. In order that any particular camera shall be capable of taking clear photographs of objects at different distances from it, it is necessary that either the focal length of the lens or its distance from the film shall be variable.

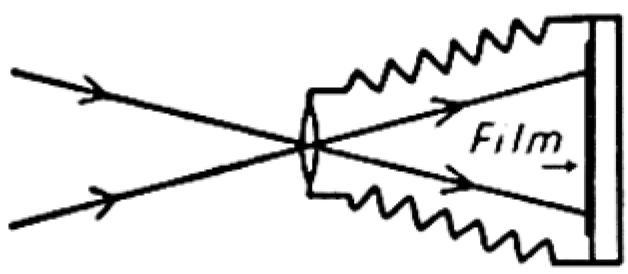


FIG. 686

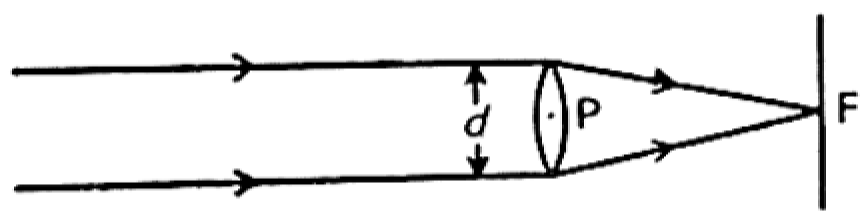


FIG. 687

If the latter adjustment is used, the sides of the camera are made of the familiar expanding bellows. These also allow the camera to be folded when not in use. The shutter, which is usually situated near the lens, cuts off the light from the film when it is closed and exposes the film for a predetermined time (*e.g.* $\frac{1}{100}$th, $\frac{1}{25}$th, 1, 5, 10 sec.), when it is opened by operating the release mechanism.

The Importance of Lens Aperture.—All the aberration problems which we have hitherto mentioned in connection with lenses are present in the design and operation of a camera. We have seen how these aberrations are, in general, reduced by stopping down the lens, and it is a fact that clear photographs can be obtained with a camera which has no very elaborately corrected lens, provided the aperture of the lens is made small. In all but the cheapest cameras the aperture can be controlled by an iris diaphragm of variable diameter. However, the quantity of light which enters the camera from any given point on the object is proportional to the effective area of the lens, and if this is made small by means of a stop there is a corresponding reduction in the quantity of light falling on the film. This necessitates a longer exposure in order to obtain a satisfactory negative, and therefore precludes the use of the camera for taking snapshots of fast moving objects.

A somewhat over-simplified explanation of the part played by the aperture of a camera lens can be made on the following lines. In Fig. 687 the lens is receiving light from a given small area of a very distant object. The image of this area will be at F, the principal focus of the lens. The

amount of light entering the lens and therefore available for forming the image of the small area considered will be proportional to the area of the lens aperture, and therefore to d^2 where d is the diameter of the aperture. Thus for lenses of the same focal length but different diameters, the illumination (*i.e.* the amount of light per unit area) of the image is proportional to d^2. On the other hand, for lenses of the same diameter the focal length affects the illumination. Remembering that the object is supposed to be infinitely distant, we can see that if the focal length is doubled, the distance from the lens to the image is doubled, so that the transverse linear dimensions of the image of the small area of the object which we are considering will be doubled. Therefore, since the lens has the same diameter as before, the amount of light entering the camera from the small area of the object is spread over an image area which is four times what it was before. Thus the illumination of the image is divided by four when the focal length is doubled, and, in general, the illumination is inversely proportional to the square of f, the focal length. Combining the effects of d and f we can say that the illumination is proportional to $\frac{d^2}{f^2}$. The number $\frac{f}{d}$, called the **stop number** or ***f*-value,** has therefore come to be used in specifying the performance of the lens in this connection. When this number is known, the time of exposure can be worked out when a film of given speed (*i.e.* light sensitiveness) is used. In any given camera (except the very simplest) the stop number can be altered by means of an adjustable diaphragm which changes the effective value of d. The setting is usually specified by expressing the value of d as a fraction of f, for example, $f/8$, $f/11$, the numeral being the stop number (f-value). Since the illumination of the image on the plate is, as we have seen, proportional to $\frac{d^2}{f^2}$, it is inversely proportional to $\left(\frac{f}{d}\right)^2$, *i.e.* to the square of stop number. Thus when a lens is stopped down from $f/8$ to $f/11$, the area of the aperture is changed by a factor $\frac{8^2}{11^2}$, *i.e.* it is approximately halved. In order to obtain the same density of negative the exposure time would therefore have to be doubled. When an actual lens is referred to as, say, $f/8$, it means that this is the largest aperture obtainable with the unstopped lens. In other words, its aperture is one-eighth of its focal length. It should be mentioned that, strictly speaking, the aperture is the *effective* rather than the actual diameter of the lens. This is known as the **entrance pupil.** In many cameras the stop numbers actually marked on the aperture adjustment are 4, 5·6, 8, 11, 16, 22, 32. Each step from one number to the next represents a halving of the area of the aperture because the squares of consecutive numbers are in the ratio of 1 : 2.

Camera Lenses.—When fast moving objects are to be photographed

it is necessary to use a short exposure time. If the photographer has no control over the lighting conditions, the short exposure necessitates a fast (*i.e.* sensitive) film and a large aperture lens (*e.g.* $f/2$). Aberrations and defects which would not be important in a lens stopped down to, say, $f/11$ can then only be avoided by using a lens in which the defects have been corrected as well as possible. Great skill and experience are required in the design and manufacture of good fast camera lenses, and the price of cameras increases sharply with the speed of the lens.

The cheapest cameras are of the fixed-focus or "box" type, in which the position of the lens is fixed relative to the film and no focusing is possible. The lens most commonly used in such cameras is the **landscape lens,** which is shown in Fig. 688. It is a converging meniscus lens mounted with the concave surface facing the incident light. At first sight this might appear contrary to the practice adopted in the case of telescope objectives, in which the convex surface faces the incident parallel light so as to distribute the deviation more uniformly between the two faces and thus to minimize spherical aberration (page 817). The requirements in the two cases are really different, however, especially with regard to the field of view, which is very much wider in a camera than in a telescope. The fact is that, by placing a stop in front of the landscape lens, those rays in an incident oblique beam which would spoil the image, *e.g.* the lowest ray in the drawing, are cut off and the remainder are brought to a sufficiently good focus on the film. The result is an image which is tolerably free from defects on account of a mutual cancellation of the various factors such as spherical aberration and coma. Curvature of the field is minimised but astigmatism is present at the edge of the image. The image is satisfactory only for apertures of $f/11$ or less, so that such a camera can only be used when illumination is good. The quality of the image is not usually good enough to bear much enlargement.

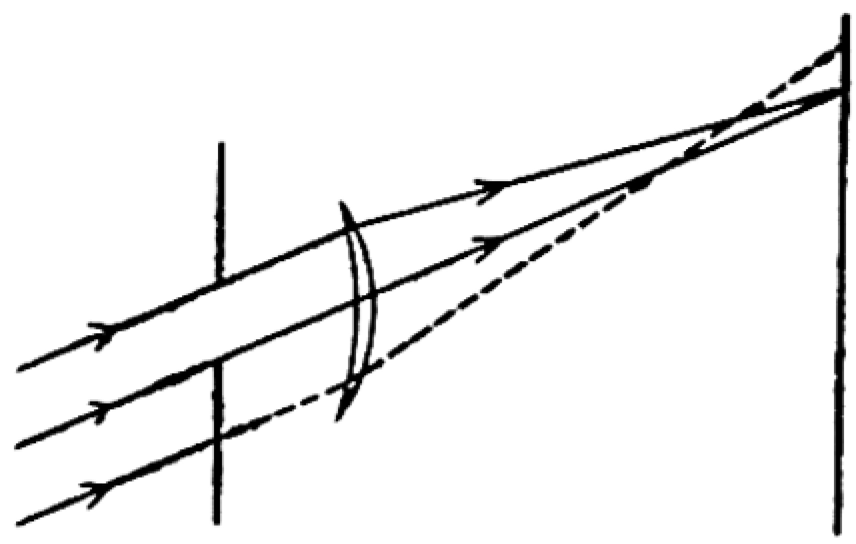


FIG. 688

It might seem at first sight that it would be impossible to obtain satisfactory photographs of objects at other than one particular distance from a camera in which the distance from the lens to the film is fixed. It is true, of course, that for a given distance between the lens and film there is only one object distance for which the image is most clearly in focus (Fig. 689 (i)). If the object is moved from this position towards or away from the lens, images of points on the object become circles as indicated in Fig. 689 (ii) and (iii). However, it is found that there is no appreciable blurring of the image unless the diameters of the circles exceed about 0·1 mm., and this fact allows a

certain latitude with regard to the object distance. This tolerance is known as **depth of field,** and its magnitude can be worked out fairly simply for any particular lens by calculating the object distance for each of the cases (ii) and (iii) in Fig. 689, assuming that the diameter of the circular patch on the film is 0·1 mm. It is evident that if the aperture of the lens is made smaller in Fig. 689 the pencil becomes narrower, with the result that for any given position of the object the circular patch on the film is smaller. Consequently the depth of field is increased by stopping down the lens. For a very small stop, conditions approach those in the pinhole camera (page 698), in which the image is in focus whatever the position of the object.

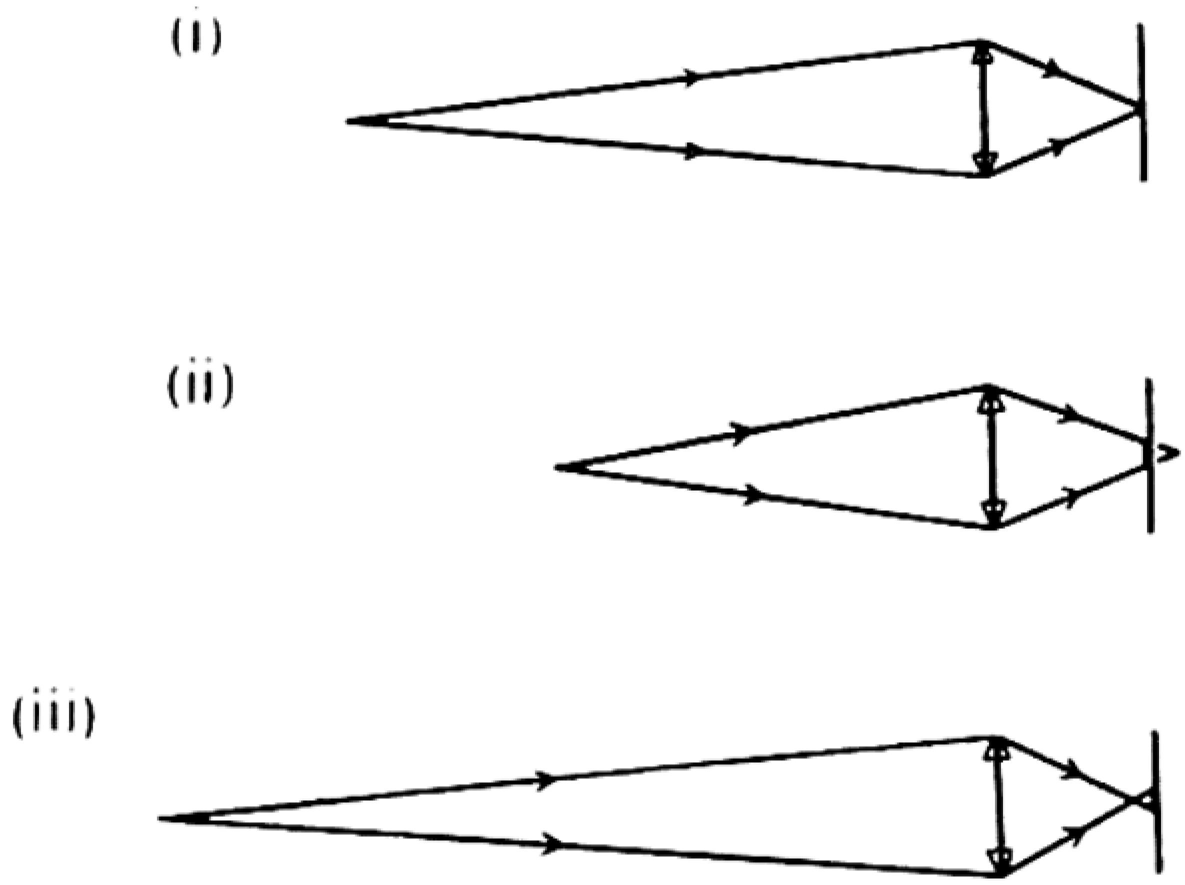


Fig. 689

In a fixed-focus camera having a landscape lens ($f/11$), objects anywhere between 6 or 8 ft. and infinity are reasonably well in focus. The question of depth of field enters into the artistic aspect of portraits and all "close-up" photography. Too great a depth of field will lead to an appearance of flatness, while too little over-enhances those features which are in focus.

The Telephoto Lens.—Everyone who has taken photographs with an ordinary camera has experienced the fact that distant objects tend to be disappointingly small on the negative. This is due to the very small ratio of image distance to object distance. The size of the image of a distant object is proportional to the focal length of the camera lens, so that to take a satisfactory photograph of a distant subject like a cricket match would require a very long and cumbersome camera if an ordinary long focal-length lens were used. The telephoto lens is designed to overcome this difficulty. It consists of a converging lens followed by a diverging lens, and its principle is illustrated in Fig. 690. Axial parallel rays from a point on a distant object are converged by the first lens towards its principal focus F′, but before reaching this point the rays are made less converging by passing through the diverging lens, and they finally come to a focus at F. The final rays are produced backwards in the drawing to cut the incident rays, and it is then clear that the system acts as a thin

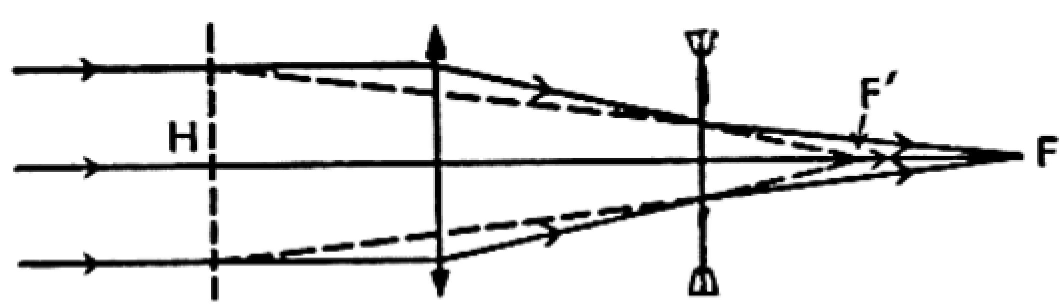


Fig. 690

converging lens with its centre at H and having a focal length HF. The point F is, of course, the second principal focus of the combination of the two lenses, H is the corresponding principal point, and HF is the focal length of the combination (page 807). Thus the effect of a long focal length is achieved without having an unduly large distance between the lens and the film.

Example.*—Find the increase in size of the image of a distant object when a diverging lens of focal length 3 in. is inserted coaxially between a converging lens, focal length 6 in., and a movable screen, the two lenses being 3·5 in. apart.* (L.I.)

When the diverging lens is inserted the system is a telephoto lens, and the ray diagram for light coming from a distant object is as in Fig. 690. The converging lens would, in the absence of the diverging lens, form a real image of the distant object at its principal focus, *i.e.* 6 in. from it. This image therefore acts as a virtual object for the diverging lens, the object distance being (6 − 3·5) in., *i.e.* 2·5 in. Thus, applying the lens equation to the diverging lens, we have

$$u = -2{\cdot}5 \text{ in.} \quad \text{(virtual object)}$$

$$f = -3 \text{ in.}$$

Therefore

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$$

$$= -\frac{1}{3} + \frac{1}{2{\cdot}5}$$

which gives

$$v = 15 \text{ in.}$$

Therefore the final real image is 15 in. beyond the diverging lens.

The magnification due to the diverging lens is equal to $\frac{15}{2{\cdot}5}$, *i.e.* 6, that is to say, the final image due to the combination is 6 times larger than the image due to the converging lens alone. This is, therefore, the answer required. It will be noticed that the combination is equivalent to a thin lens of focal length 6 times that of the converging lens, *i.e.* a focal length of 36 in.

5. PROJECTORS

Of all the other optical instruments which we have not so far described, we shall deal only with projectors. The first of these is the optical lantern, known for many years as the "magic lantern," which is used for the projection of still transparencies of comparatively large size such as lantern slides.

The Optical Lantern.—In this instrument light passes through a transparent lantern slide on which is printed a positive photograph or drawing, and this light is then used by the projection lens to form a greatly magnified real image of the slide on a screen. The arrangement is shown in Fig. 691. In order to exhibit the formation on the screen of images of points on the slide, it is necessary to show two rays from a point such as O proceeding through the projection lens and thereby being converged to the

image of O at I. At the other end of the slide O′ gives rise to an image I′. The rays travelling from O to the projection lens come, in the first place, from the source of light, and, since we have drawn two such rays, they must initially have come from two different points on the source, namely S_1 and S_2. They cannot come from one and the same point, otherwise they would be indistinguishable all the way through the system. The rays which go to form the image I′ of O′ are also shown as originating at S_1 and S_2. In fact, of course the light reaching I comes from all points

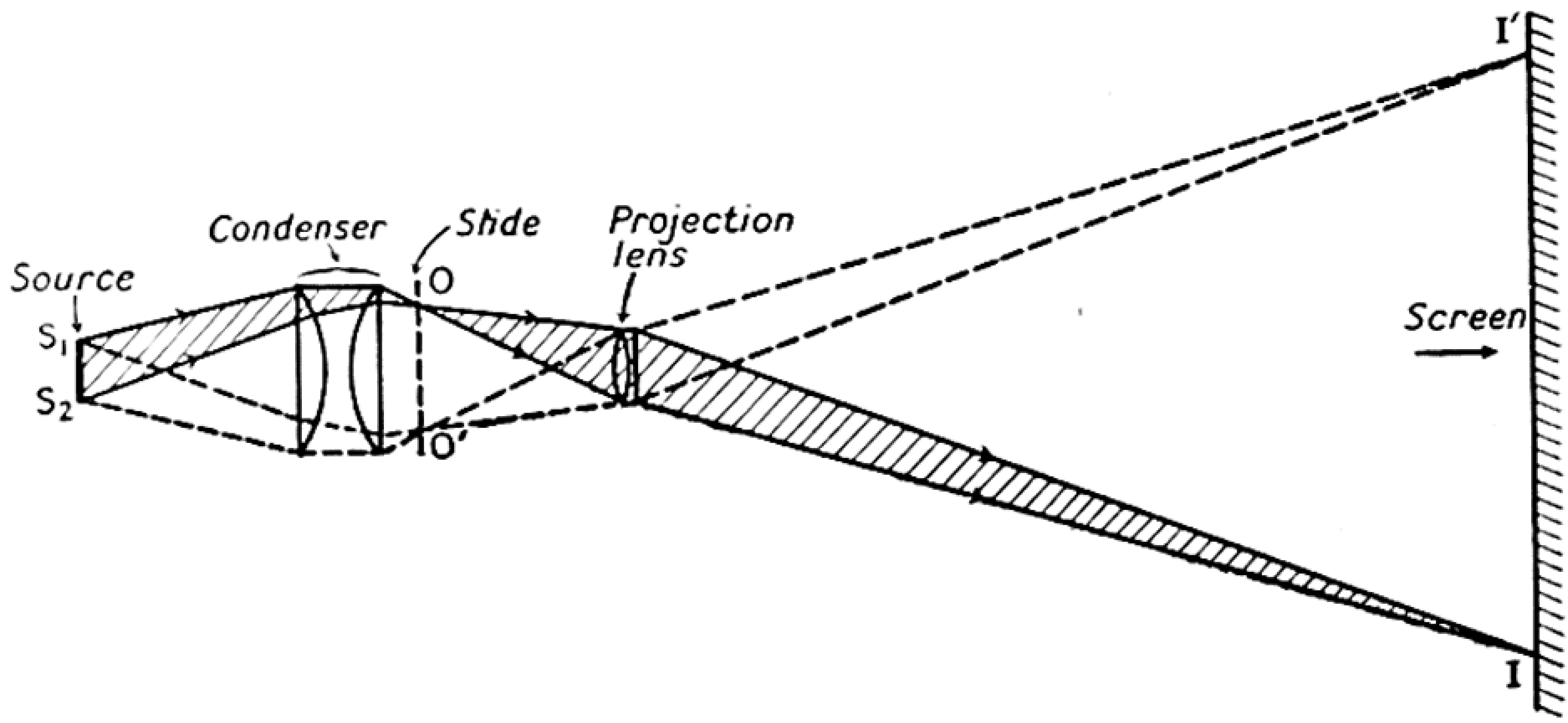


FIG. 691

between S_1 and S_2, and the pencil bounded by the extreme rays is shaded to indicate this fact. The images of all the points on the slide between O and O′ are formed by similar pencils bounded by lines originating from S_1 and S_2.

The source of light is frequently the crater of an electric carbon arc, and the condenser lens often consists of two separate plano-convex lenses with the plane faces outwards so as to distribute the deviation and minimize spherical aberration. The action of the condenser, which must obviously be at least as wide as the slide, is to converge the light passing through the slide so that all of it passes through the projection lens. The latter is placed at or near the narrowest part of the beam (where in fact the condenser forms a real image of the source) so as to allow all the available light to pass through it. It is evident that in the absence of the condenser only that central portion of the slide through which light would pass on its rectilinear path from the source to the projection lens would be projected on the screen.

The projection lens must be as free from chromatic and other aberrations as possible, and of such a focal length as to enable it to give a focused image of the slide when placed at approximately the narrowest part of the beam given by the condenser. Focusing is carried out by moving the projection lens. The transverse magnification of the image is, of course,

equal to the ratio of the distances of the projection lens from the screen and from the slide.

Cinematograph Projector.—The "frame" of a cinematograph film is considerably smaller in size than a lantern slide. If the same system as that shown in Fig. 691 for lantern-slide projection were used for the film, the condenser would be smaller and the amount of light collected and projected by it would be correspondingly reduced. A different system is therefore adopted, as shown in Fig. 692, in which the light from the

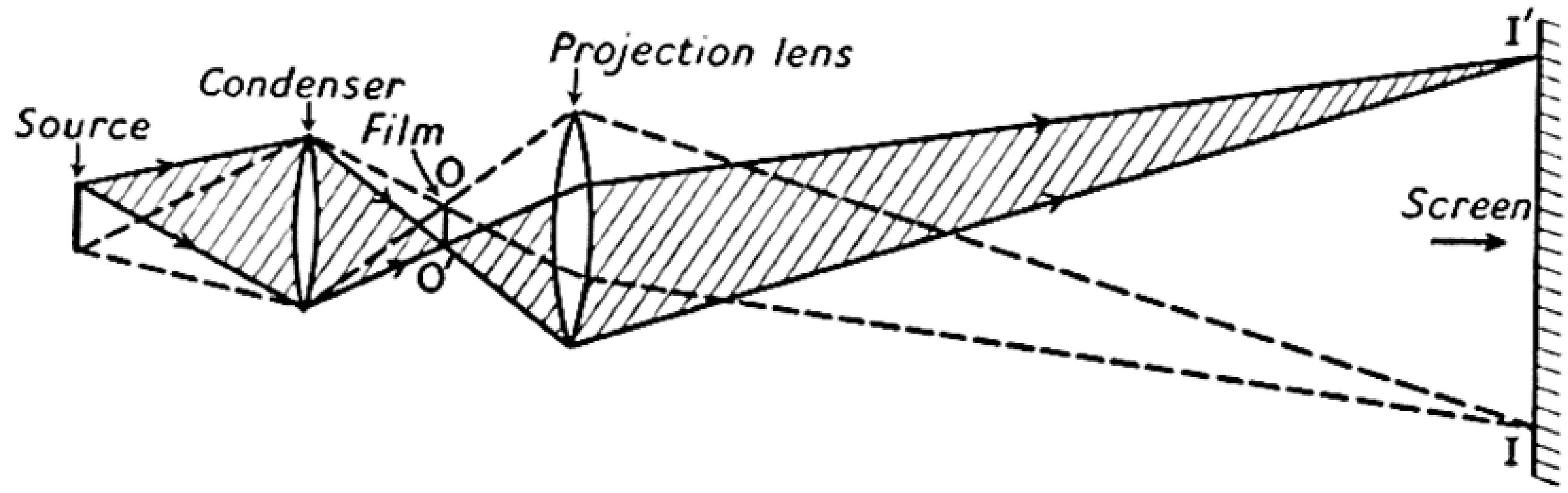


FIG. 692

condenser is concentrated through the film by placing the latter at the image of the source formed by the condenser. In practice the condenser lens is frequently replaced by a concave reflector, placed behind the source, which has the same effect of forming a real image of the source in the plane of the film.

EXAMPLES XLIX

1. Define magnification and magnifying power. Calculate the magnification produced by a converging lens of 1 cm. focal length when used as a "magnifying glass" by a person whose near point is 25 cm. in front of his eye. Why is there very little noticeable chromatic aberration when a single lens is used in this way? (C.H.S.)

2. Draw a diagram to show how two lenses can be arranged to form a microscope. What are the principal defects in the image produced by such a microscope? Indicate how *one* of these defects is overcome in practice.
Why are oil-immersion objectives used for high-power work? (L.Med.)

3. Explain the principle of a microscope consisting of two simple convex lenses, and illustrate your answer by a diagram showing the passage through the instrument of a pencil of rays from a non-axial point on the object. Explain as fully as you can why, in practice, the objective of a microscope consists of a number of lenses, and the eyepiece generally contains two lenses not in contact. (L.Med.)

4. Explain the use of a reading lens, and establish a formula for the magnification when used, close to the eye, by an observer whose least distance of distinct vision is d.
A concave lens of 5 cm. focal length is placed 11 cm. in front of a convex lens of 10 cm. focal length. Find the magnification and position of the image of an object placed 20 cm. in front of the concave lens. (L.I.)

5. Explain, using diagrams, the action of a simple magnifying-glass and deduce the magnifying power in terms of f and D, the least distance of distinct vision.

Draw the path of two rays through a microscope which uses two lenses, showing the formation of the magnified image, and explain the meaning of magnifying power in this case. Indicate the method you would use to measure this. (L.H.S.)

6. Draw a ray diagram to illustrate the formation of the image of a non-axial point on an object by two thin converging lenses arranged to act as a microscope.

If the focal lengths of objective and eye-lens are 1 inch and 2 inches respectively and their distance apart is 6 inches, where must an object be placed in order that the image seen by the eye may be 12 inches from the eye-lens? What is the magnification produced? (L.I.)

7. Explain the use of a converging lens as a simple microscope and obtain an expression for the magnifying power.

A model compound microscope consists of two thin converging lenses, an objective of 1 cm. focal length and an eyepiece of 3 cm. focal length separated by 24 cm. Find where the object must be placed so that the final image may be at infinity. Give ray diagrams. (Assume nearest distance of distinct vision = 25 cm.) (L.I.)

8. A microscope is first focused on an object on a slide, and a glass cover-slip, 0·6 mm. thick and of refractive index 1·5, is then placed over the object. Through what distance and in what direction must the microscope be moved to restore focus?

Draw a diagram showing how two lenses, of powers 20 and 10 dioptres respectively, may be arranged to demonstrate the action of a microscope. Indicate the positions of the axial points of the object and first image relative to the first focal points of the lenses, and draw the courses of two rays from a non-axial point on the object to the eye of the observer. (L.Med.)

9. Describe how two thin lenses may be set up to form a compound microscope, and trace right through the system the paths of a pencil of rays originating from a point on the object that is not on the axis of the microscope.

Mark on this diagram the best position for the user's eye.

Two simple lenses, an objective of focal length 1 cm. and an eyepiece of focal length 5 cm., are set up 18 cm. apart, and arranged to give a final image 25 cm. from the eyepiece. Find (*a*) the distance of the object from the objective, (*b*) the magnification obtained. (O.H.S.)

10. What do you understand by (*a*) the apparent size of an object, and (*b*) the magnifying power of a microscope?

A model of a compound microscope is made up of two converging lenses of 3 and 9 cm. focal length at a fixed separation of 24 cm. Where must the object be placed so that the final image may be at infinity? What will be the magnifying power if the microscope as thus arranged is used by a person whose nearest distance of distinct vision is 25 cm.? State what is the best position for the observer's eye, and explain why. (L.H.S.)

11. Describe the optical arrangement of a compound microscope, and give a diagram showing the path of a pencil of rays from a non-axial point on the object through the instrument to the eye of an observer.

An object lens of 2·0 cm. focal length and an eye-lens of 6·0 cm. focal length are placed coaxially 16 cm. apart, and adjusted so that the virtual image seen through the eye-lens is 30 cm. from that lens. Determine the magnification obtained. (L.I.)

12. Explain what is meant by the magnifying power of an optical instrument, considering the cases of microscope and telescope.

A thin converging lens of focal length 5 cm. is laid on a map situated 60·5 cm. below the eye of an observer whose least distance of distinct vision is 24·5 cm. Describe what is seen (*a*) then, and when the lens is raised, (*b*) 5 cm., (*c*) $5\frac{1}{2}$ cm., (*d*) 6 cm. above the map. (L.H.S.)

13. Define *angular magnification* of a telescope.

A telescope made up of two convex lenses 25 cm. and 5 cm. focal length is focused so that the image of a scale 150 cm. from the object glass coincides with the scale itself. Determine the angular magnification produced. (L.I.)

14. Describe the optical arrangement of an astronomical telescope, illustrating your description with a diagram.

A simple astronomical telescope has two lenses of 60 cm. and 5 cm. focal length respectively. It is used to view a distant object by an observer whose least distance of distinct vision is 25 cm. What is the angular magnification obtained when the image is formed (*a*) at infinity, (*b*) at the observer's least distance of distinct vision?

15. Explain what is meant by the magnification (or magnifying power) of an astronomical telescope, and describe how you would attempt to measure it for a given instrument set up in normal adjustment.

A small astronomical telescope has an objective of focal length 50 cm. and an eyepiece of focal length 5 cm. The lenses are separated to a distance of 56 cm. and the instrument is directed at the sun. Find the position, nature and size of the final image if the angle subtended by the sun at a point on the earth's surface is 32′. (O.H.S.)

16. Describe, with the aid of a diagram, the construction and mode of action of an astronomical telescope and obtain an expression for its magnifying power. What advantages are to be gained by using a concave mirror as an object glass?

A telescope is constructed of two converging lenses of focal lengths 100 cm. and 20 cm. placed so that the final image is at infinity. Calculate (i) the overall magnification, and (ii) the size of the image formed by the object glass of a building 60 metres high and distant one kilometre. (L.A.)

17. Explain the essential features of the astronomical telescope.

Define and deduce an expression for the magnifying power of this instrument.

A telescope is made of an object glass of focal length 20 cm. and an eyepiece of 5 cm., both converging lenses. Find the magnifying power in accordance with your definition in the following cases: (*a*) when the eye is focused to receive parallel rays, and (*b*) when the eye sees the image situated at the nearest distance of distinct vision, which may be taken as 25 cm. (L.I.)

18. Draw a diagram of the Galilean telescope or opera glass.

A converging lens of focal length 12 cm. and a diverging lens of focal length $2\frac{1}{4}$ cm. are situated 10 cm. apart. An object is situated 3 m. from the converging lens. Find the position of the final image produced by refraction, first by the converging lens and then by the diverging lens.

What is the ratio of the angle subtended at the centre of the diverging lens by the final image to the angle subtended by a small object? (L.I.)

19. Describe how to determine the focal length of a diverging lens using an auxiliary converging lens not placed in contact with it.

An opera glass contains a converging lens of 4 in. focal length and a diverging lens of 1·5 in. focal length. What is the separation of the two lenses when a virtual image of a distant object is formed at a distance of 10 in. from the eye-lens, and what is then the magnifying power? Before commencing your calculation state the sign convention you will use. (L.I.)

20. A convex lens of focal length 20 cm. and a concave lens of focal length 10 cm. are arranged for use as an opera glass. Draw a ray diagram to scale showing how the final image at infinity is produced, describing briefly how you do this, and derive the magnifying power.

When an object is placed 60 cm. in front of the convex lens and the lenses are separated by a distance x, a real image is formed 30 cm. beyond the concave lens. Calculate x. (C.H.S.)

21. Show by means of a diagram that, when a simple converging lens held close to the eye is used as a magnifying-glass, it is achromatic in the sense that the various differently coloured images all subtend the same angle at the eye.

What is meant by saying that a telescope objective is achromatic, and how is this condition achieved? (L.Med.)

22. Explain briefly what is meant by rectilinear propagation of light and describe the refraction of light by a lens.

Using illustrative ray diagrams, compare the action of a pinhole camera with that of a fixed-focus lens camera. What is meant by depth of focus? On what factors do the definition and the brilliancy of the resulting images depend? (L.Med.)

Chapter L

PHOTOMETRY

1. PRINCIPLES AND DEFINITIONS

Photometry concerns the measurement of quantity of light. It has been explained on pages 527–530 (Vol. 2) that the radiation given out by a heated body is of the same physical character whether the temperature is high enough to render the body visible by means of the radiation or not. Radiation within a certain wave-length range is visible to the eye, but its physical nature is the same as the radiation outside this range. Radiation can be measured by its heat energy irrespective of any effect it may have on the eye, but when we are concerned with the measurement of the illuminating powers of lamps or the illumination of objects in a room or on a public highway, we must make our definitions and measurements with reference to the visual effect of radiations rather than to their actual physical energies. In this connection we may note two facts. Firstly, the apparent brightness (*i.e.* the intensity of the sensation produced in the eye) of visible radiations of the same energy but different wave-lengths depends on the wave-length (or colour) of the light. In other words, the eye is not equally sensitive to all visible wave-lengths. Secondly, the sensation of brightness produced by radiation of a given wave-length is not proportional to the energy of the radiation. For moderate amounts of light the Weber-Fechner law referred to in connection with the loudness of sounds (page 629, Vol. 3) is fairly closely obeyed, and this means that equal increments of apparent brightness are produced by equal *fractional* (or *percentage*) increases of the energy received by the eye. That is to say, the response or sensation is proportional to the logarithm of the stimulus measured in energy units.

Luminous Flux. Luminous Intensity.—As a first step towards putting the question of light measurement on a quantitative basis we state that by the term "amount of light" we mean the "amount of radiant energy evaluated by reference to the luminous sensation produced." This may seem rather vague at first. We shall see, however, that this and other definitions can be used in conjunction with a system of measurement based on visual perception.

The first definition is that of the **luminous flux** passing through a given area or leaving a given source. This stands for the amount of light (in the sense indicated above) flowing in unit time. Clearly we can express the light-giving power of a source of light by stating the luminous

flux which it emits, but since we have not yet established an experimental method of measuring luminous flux—or indeed a unit—it is necessary to introduce and define some sort of reproducible standard source of light, the luminous flux from which can be taken as the basis of the definition of the unit of this quantity. In earlier times the flame of a candle of specified dimensions, composition and rate of burning was taken as the standard light source. This was later replaced by the Vernon Harcourt lamp which was of specified construction and burnt pentane under definite stated conditions. Since 1948 it has been agreed internationally that the standard source of light is 1 cm.^{2} of a black body maintained at the melting-point of platinum. The meaning and method of reproduction of a black body are explained on pages 534–535 (Vol. 2), and we need only say here that the definition refers to the radiation which is emitted per sq. cm. from a very small hole in the wall of an enclosure maintained at the stated temperature. A uniform *point* source of light (if such could exist) from which the total luminous flux is one-sixtieth of that emitted by the standard is said to have a **luminous intensity** of one **international candle,** and the luminous intensities of other sources which can be treated as points are defined as being in proportion to the total luminous flux which they emit. Other terms synonymous with luminous intensity are **illuminating power** and **candle power.**

The international standard light source is not, of course, reproduced every time the luminous intensity of a source is measured. Use is made of secondary standards in the form of electric lamps which have been calibrated directly or indirectly against the standard.

Obviously we can define the unit of luminous flux by reference to the emission of light from a source of unit candle power, and this is done as follows. Suppose that a point source having a luminous intensity or candle power of one international candle is situated at O (Fig. 693) and that we imagine a sphere of radius r and centre O to be described round the source. The light energy emitted by the source travels along the radii of the sphere and passes normally through the surface of the sphere. The area of the surface of the sphere is $4\pi r^2$, and if a cone with apex at O cuts off an area A of the surface, then, provided that the emission from O is uniform in all directions, the luminous flux emitted into this cone will be $\dfrac{A}{4\pi r^2}$ of the total flux emitted. The quotient $\dfrac{A}{r^2}$ is called the **solid angle** of the cone, or the solid angle subtended by the area A at the centre of the sphere. The solid angle is unity when A is equal to r^2, and the solid angle corresponding to the whole sphere is 4π. The unit of luminous flux—known as the **lumen**—is defined as the flux emitted by a point source of luminous intensity 1 c.p. into a cone of unit solid angle, or

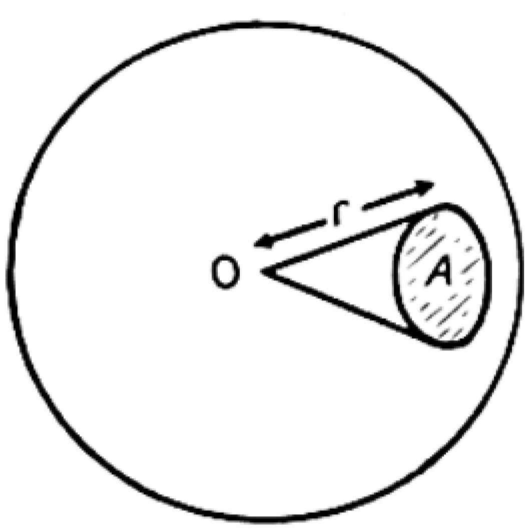


FIG. 693

alternatively, the flux falling on unit area of the surface of a sphere of unit radius when a source of 1 c.p. is placed at its centre. The lumen is therefore $\frac{1}{4\pi}$ of the total luminous flux emitted by a uniform point source of 1 c.p. In other words, a source of luminous intensity 1 c.p. emits 4π lumens and, since the luminous intensity of a source is, by definition, proportional to the flux which it emits, a source has an intensity of I c.p. if its total emission is $4\pi I$ lumens. Furthermore, the flux emitted into a solid angle ω by a source of I c.p. is $I\omega$.

The Inverse Square Law.—Suppose that a point source O (Fig. 694 (i)) has an intensity of I c.p. and that a small area ΔA is situated at a distance

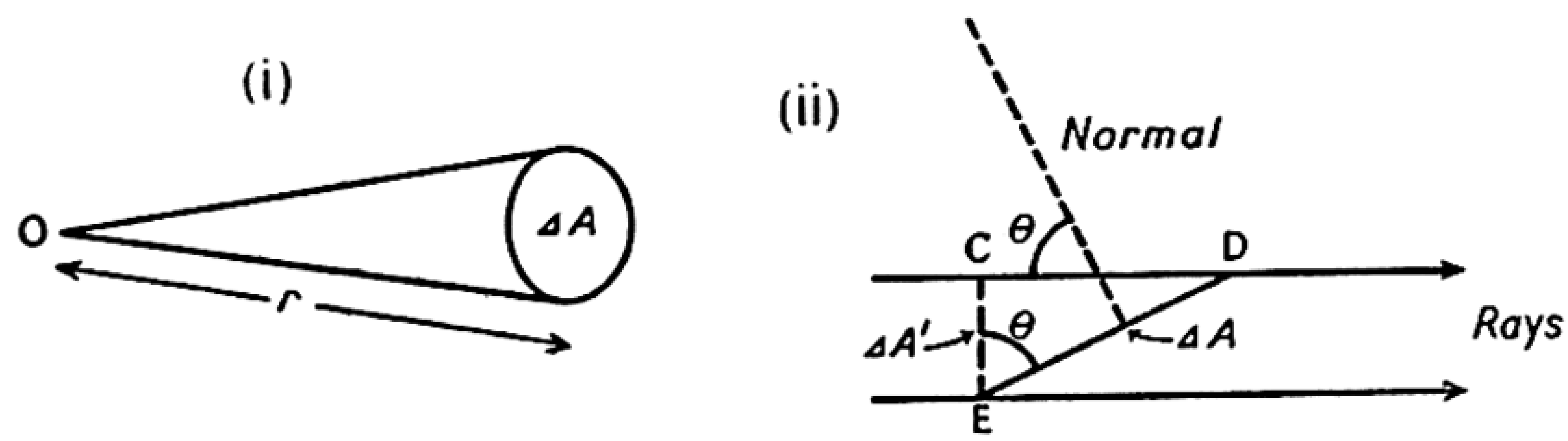


FIG. 694

r from O and is perpendicular to r. The small solid angle subtended by ΔA at O is $\frac{\Delta A}{r^2}$, so that the flux ΔF emitted by O into this solid angle is given by

$$\Delta F = I\frac{\Delta A}{r^2}$$

The flux ΔF falls on ΔA, and the fact that ΔF is proportional to $\frac{1}{r^2}$ for a given area placed perpendicular to the rays from a given point source is known as the **inverse square law.** Its truth depends on the rectilinear propagation of light and on the absence of absorption in the medium between O and ΔA.

The Cosine Law.—Now suppose that ΔA is sufficiently small and that r is sufficiently large to allow us to regard the rays from O as being sensibly parallel to each other in the neighbourhood of ΔA. Furthermore, let the normal to ΔA be no longer parallel to the rays but inclined to them at an angle θ as in Fig. 694 (ii). The line CE drawn perpendicular to the rays represents the projection of ΔA on a plane perpendicular to the rays. Let the area of this projection be $\Delta A'$. It is easily seen that $\widehat{CED} = \theta$, so that $CE = DE \cos\theta$. Thus if ΔA is divided into strips whose lengths are parallel to the plane of the paper, the length of each is multiplied by $\cos\theta$ when it is projected on to $\Delta A'$ but its width is unaltered. The area of each strip in $\Delta A'$ is therefore $\cos\theta$ times the corresponding strip

in ΔA, so that if we add up all the strips in each area we have

$$\Delta A' = \Delta A . \cos \theta$$

In Fig. 694 (ii) the flux ΔF through ΔA is evidently the same as that through $\Delta A'$ and, by the inverse square law, is given by

$$\Delta F = I \frac{\Delta A'}{r^2}$$

$$= I \frac{\Delta A . \cos \theta}{r^2} \quad . \quad . \quad . \quad . \quad (1)$$

Illumination.—The illumination (sometimes called "intensity of illumination" and also **illuminance**) of a surface is defined as the luminous flux incident per unit area on the surface. In other words, the illumination at a point on a surface is equal to $\frac{\Delta F}{\Delta A}$, where ΔF is the flux falling on a small area ΔA surrounding the point.

Reference to expression (1) shows that the illumination of a small area ΔA at a distance d (using this symbol instead of r to denote distance) from a point source of intensity I c.p., the normal to ΔA making an angle θ with the direction of the light, is equal to

$$\frac{I \cos \theta}{d^2} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The value of this expression will be unity if the area is placed at right angles to the line joining it to the source ($\theta = 0$, $\cos \theta = 1$) and if I is made 1 c.p. and d is made unit distance. If the unit chosen for d is the foot, then the unit of illumination is the illumination on a screen placed at 1 ft. from a source of 1 c.p. and perpendicular to the rays. This unit is the **foot-candle.** For a screen of finite size the illumination in such a case will not be uniform all over the square foot of screen unless the screen is part of the surface of a sphere with the source at the centre. An area of 1 ft.^{2} of such a spherical screen would subtend unit solid angle at the source and would therefore receive a luminous flux of 1 lumen. Therefore a screen on which the illumination is 1 ft.-candle receives 1 lumen ft.$^{-2}$. In fact this is a rather more logical method of expressing the unit of illumination, because "ft.-candle" does not stand for the product of length and candle power like "ft.-poundal" does for length and force in the unit of work.

If the metre is adopted as the length unit, then the unit of illumination becomes the **metre-candle,** which is the illumination on a screen situated at right angles to the light at a distance of 1 m. from a point source of 1 c.p. Each square metre of such a screen (supposing it to be spherical and of radius 1 m.) will receive a luminous flux of 1 lumen, so that the

metre-candle is the same unit as the lumen $m.^{-2}$. It is also given the separate name **lux.**

Fig. 695 represents a plane surface with a point source of light situated in front of it at O. The line OP is perpendicular to the plane while OQ makes an angle θ with the normal to the plane at Q. If the source at O gives out light uniformly in all directions and has an intensity of I c.p., then, according to expression (2), the illumination at P (E_0) will be given by

$$E_0 = \frac{I}{(OP)^2}$$

since $\cos\theta = 1$ in this case. The units of E_0 will be ft.-candles if OP is

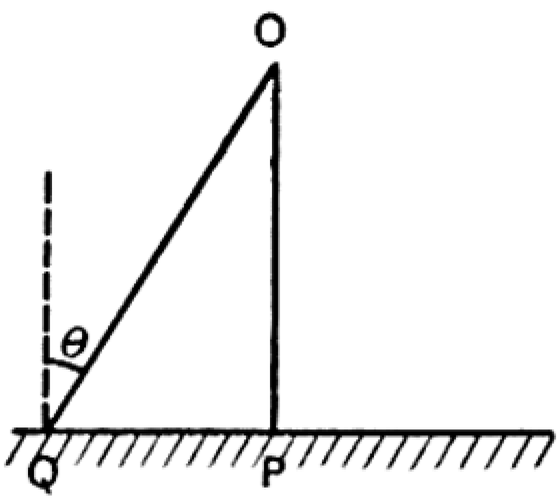


FIG. 695

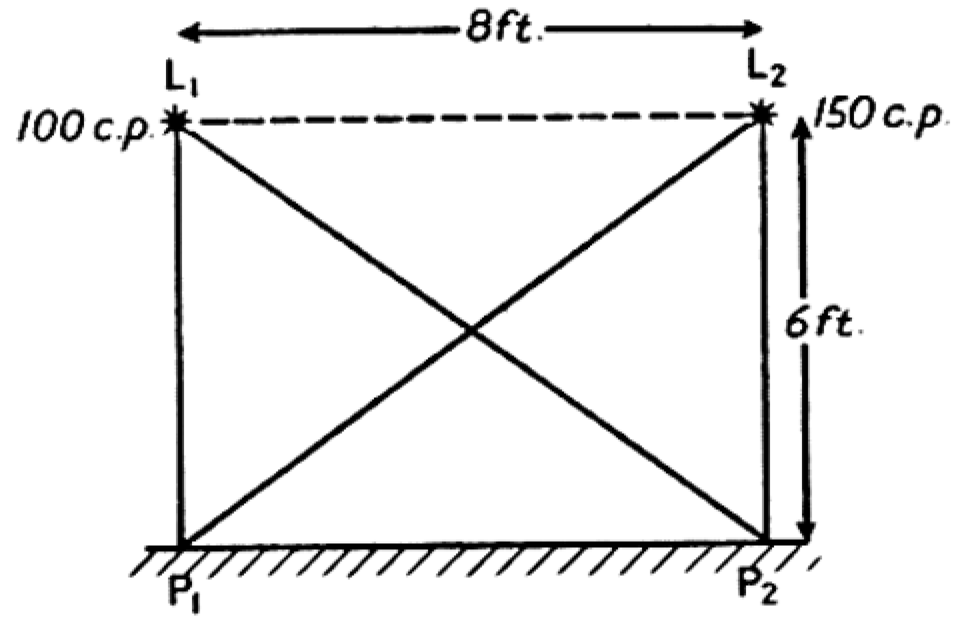


FIG. 696

in feet, or metre-candles (lux) if OP is in metres. At Q the illumination (E_θ) will be given by

$$E_\theta = \frac{I\cos\theta}{(OQ)^2}$$

$$= E_0\left(\frac{OP}{OQ}\right)^2 \cos\theta$$

$$= E_0 \cos^3\theta$$

since $\cos\theta = \cos P\widehat{O}Q = \frac{OP}{OQ}$. The fact that in going outwards along the surface from P the illumination is proportional to $\cos^3\theta$ is sometimes referred to as the "(cosine)^3 law."

***Example.**—Two lamps, of candle-powers* 100 *and* 150 *respectively, are suspended* 6 *ft. above a horizontal surface with a distance of* 8 *ft. separating them from each other. Calculate the illumination vertically below each lamp.* (L.I.)

At P_1 (Fig. 696) vertically below L_1 (the 100 c.p. lamp) the illumination due to L_1 is $\frac{100}{6^2}$ ft.-candles, while that due to L_2 is equal to $\frac{150}{(L_2P_1)^2} \times \cos L_2\widehat{P_1}L_1$ ft.-candles.

Inspection of triangle $L_2P_1L_1$ shows that $L_2P_1 = 10$ ft., and therefore $\cos L_2\widehat{P_1}L_1$ is equal to $\frac{6}{10}$, so that the total illumination at P_1 is equal to

$$\frac{100}{36} + \frac{150 \times 6}{1000} \text{ ft.-candles}$$

which is equal to

$$2{\cdot}78 + 0{\cdot}90 \text{ ft.-candles}$$

or

$$3{\cdot}68 \text{ ft.-candles}$$

At P_2 the illumination is equal to

$$\frac{150}{6^2} + \left(\frac{100}{10^2} \times \frac{6}{10}\right) \text{ ft.-candles}$$

which is equal to

$$4{\cdot}17 + 0{\cdot}6 \text{ ft.-candles}$$

$$4{\cdot}77 \text{ ft.-candles}$$

Brightness.—It is important to realize that in photometry the term illumination refers to the light *falling on* the surface. How bright the surface may appear to be as a result of a given illumination depends upon the nature of the surface itself, *e.g.* upon its colour and texture. In general, too, we must expect the brightness to depend upon the angle at which the surface is viewed.

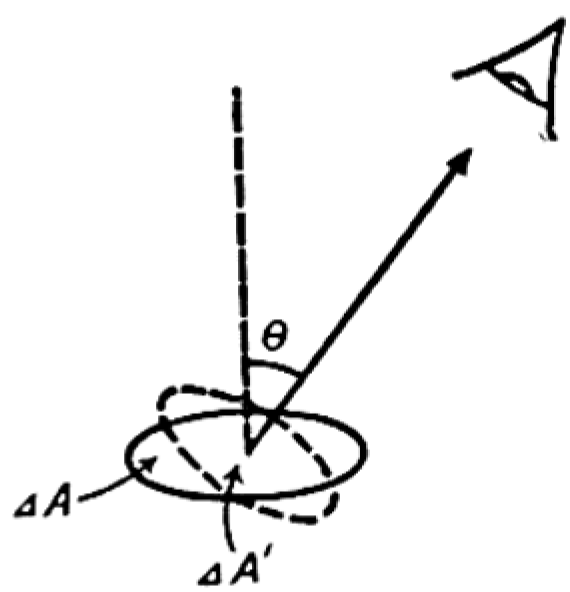


FIG. 697

Suppose that a small area ΔA of a surface (Fig. 697) has an equivalent candle-power of ΔI_θ in a direction making an angle θ with the normal to it. This means that the surface emits a flux of $\Delta I_\theta . \Delta\omega$ lumens into a small solid angle $\Delta\omega$ in the given direction. Since we are concerned with the brightness of the surface, which is a quantity depending, so to speak, on the "density" of the light rather than the area of the emitting surface, we express the brightness as the equivalent candle-power *per unit effective* area. To an eye placed as in Fig. 697 to receive the oblique emission, the effective area of ΔA is $\Delta A'$, the projection of ΔA on to a plane perpendicular to the direction of emission considered. As explained on page 902,

$$\Delta A' = \Delta A . \cos\theta$$

The brightness (B_θ) of ΔA in the given direction is accordingly given by

$$B_\theta = \frac{\Delta I_\theta}{\Delta A'}$$

$$= \frac{\Delta I_\theta}{\Delta A . \cos\theta} \qquad . \quad . \quad . \quad . \qquad (3)$$

Brightness is evidently expressed in candles per unit area. An alternative name for it is **luminance.**

When light falls on a rough material with no regular structure, such as blotting paper or plaster of paris, it is reflected in all directions owing to the unevenness of the surface, and the result is that the brightness of the surface appears to vary very little with the angle from which it is viewed. This fact has led to the conception of an ideal surface which, when uniformly illuminated, appears equally bright from all directions. It is called a "perfectly diffusing surface." For such a surface the quantity B_θ in equation (3) can be replaced by the constant brightness B, and the equation can then be written

$$\begin{aligned} \Delta I_\theta &= B.\Delta A.\cos\theta \\ &= \Delta I_0.\cos\theta \end{aligned}$$

where $\Delta I_0 (= B.\Delta A)$ is the candle-power of the surface ΔA in the direction of the normal. Therefore the candle-power of the perfectly diffusing surface in a direction making an angle θ with the normal is proportional to $\cos\theta$. This relationship is known as **Lambert's law.**

When an extended source of light or a diffusing surface is viewed by the eye, it appears equally bright *at all distances* provided that the diameter of the pupil of the eye remains unaltered. This is so because the illumination of the image on the retina of the eye is the determining factor. As the distance between the source and the eye is increased, the flux entering the pupil diminishes according to the inverse square law, but the area of the retinal image also diminishes in the same way so that the illumination of this image remains constant.

2. THE DETERMINATION OF CANDLE-POWER

The Principle.—Suppose that a source of light L_1 which is small enough to be regarded as a point source, and which has an intensity of I_1 candles, is situated in front of a small screen with a rough diffusing surface, the perpendicular distance from the source to the screen being d_1 (Fig. 698). The illumination of the screen at the foot of the perpendicular is $\frac{I_1}{d_1^2}$. This, it will be remembered, is the flux falling on unit area of the surface. The flux diffused from unit area of the surface will be proportional to the incident flux and can be written as $\frac{r_1 I_1}{d_1^2}$, where r_1 is a factor depending on the nature of the surface and called the reflecting factor. The brightness of the surface as defined on page 904 (*i.e.* the flux emitted per unit area per unit solid angle) will evidently be proportional to the total flux emitted by unit area, so that we can take the latter as a measure of the brightness. (For a perfect diffuser the total flux emitted can be shown to be equal to π times the brightness.)

Fig. 698 shows a similar surface to the first placed near to it and illuminated by a small source L_2 of candle-power I_2 at a distance d_2. We must imagine that an opaque non-reflecting partition is placed so as to ensure that each surface is illuminated by only one source. The illumination of the second surface is equal to $\frac{I_2}{d_2^2}$, and the flux diffused by unit area will be equal to $\frac{r_2 I_2}{d_2^2}$, where r_2 is the reflecting factor for the second surface.

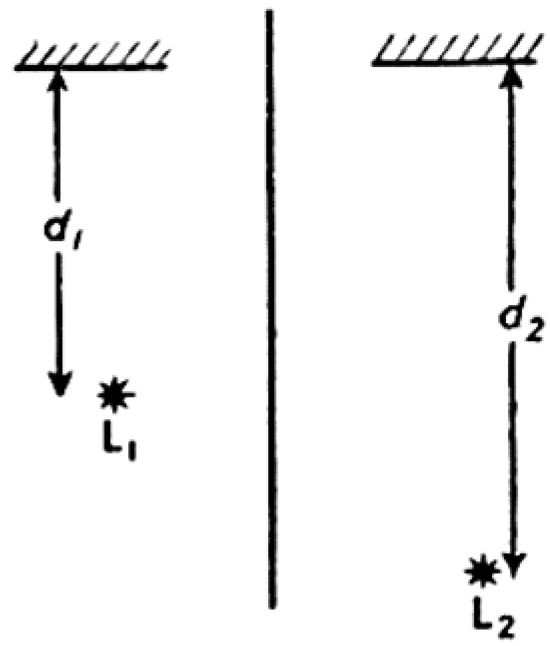


Fig. 698

The eye is not able to judge the relative brightness of two surfaces of unequal brightness, but it can judge equality of brightness, and we base a system of visual photometry on this principle. Since r_1 and r_2 are usually unknown, the simplest plan is to make the two surfaces as similar as possible and then to assume that $r_1 = r_2$. The distances of the sources are then altered until the two surfaces appear equally bright, and when this condition is achieved the illumination of each surface is the same, so that

$$\frac{I_1}{d_1^2} = \frac{I_2}{d_2^2}$$

or

$$\frac{I_1}{I_2} = \frac{d_1^2}{d_2^2} \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The ratio of the intensities of the two sources can therefore be calculated when d_1 and d_2 have been measured, and if one of the intensities is known, the other can be found. Such an arrangement as that shown in Fig. 698 does not constitute an accurate photometer, but it serves to explain the principle by which the judgment of equal brightness by the eye is used in photometry.

In photometers based on the foregoing principle it may well be that r_1 and r_2 are not exactly equal. In this case the true equation is

$$\frac{r_1 I_1}{d_1^2} = \frac{r_2 I_2}{d_2^2} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

and the result would be in error if we used equation (4). Suppose, however, that we now interchange the two sources so that each illuminates the surface previously illuminated by the other and that we again match the brightness. If the distance of L_1 from the surface it now illuminates is d_1' and the distance of L_2 is d_2' we have

$$\frac{r_1 I_2}{d_2'^2} = \frac{r_2 I_1}{d_1'^2} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Dividing equation (5) by equation (6) we obtain

$$\frac{I_1}{I_2}\cdot\frac{d_2'^2}{d_1^2}=\frac{I_2}{I_1}\cdot\frac{d_1'^2}{d_2^2}$$

so that

$$\frac{I_1}{I_2}=\frac{d_1d_1'}{d_2d_2'} \qquad . \quad . \quad . \quad . \quad . \quad (7)$$

Hence, to avoid errors due to the inequality of r_1 and r_2 in any photometer, the interchange should be made and equation (7) used.

One very important condition which must be fulfilled before the judging of equal brightness is even possible is that the colours of the two lights must be the same. Not only is it impossible to do the matching with, say, red and blue lights, but it is also difficult and uncertain if the two sources are, for example, electric filament lamps through which the currents are such as to produce different filament temperatures and therefore different qualities of radiation.

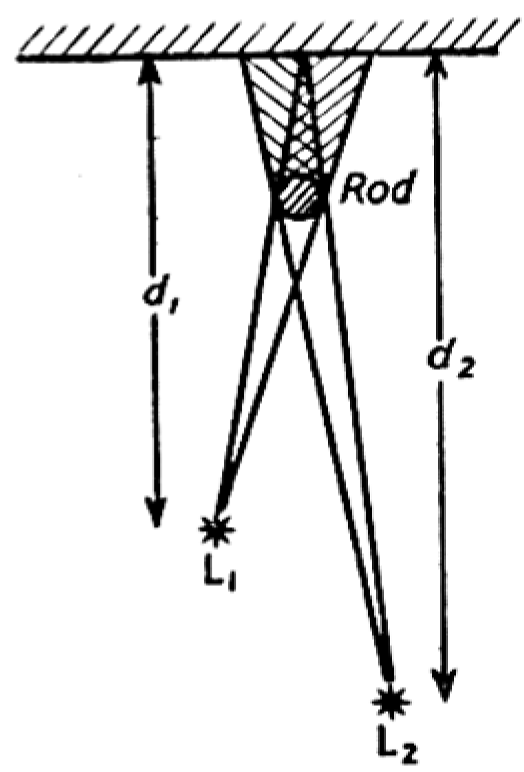


FIG. 699

Rumford's Shadow Photometer.—This is a very simple device for achieving the conditions described above. Fig. 699 is a plan of a simplified version of Rumford's original arrangement. A vertical rod is set up in front of a screen of rough white paper and the two sources L_1 and L_2 (which must be small in size) are placed so as to produce shadows of the rod on the screen. The positions of the sources are adjusted until the shadows are side by side and are judged to be of equal brightness. The shadow cast by either one of the sources is illuminated by the other, so that when the adjustment of equal brightness has been made we know that the illumination produced on the screen by each source is the same. Therefore, if I_1 and I_2 are their candle-powers and d_1 and d_2 are their respective distances *from the screen*, we know that

$$\frac{I_1}{I_2}=\frac{d_1^2}{d_2^2}$$

It is interesting to note that a uniform general illumination of the screen by some other source would not vitiate the result. Furthermore, in the absence of such extraneous light, the brightness of the screen just outside the region of the shadows is double that inside the shadows.

Joly's Block Photometer.—This is a very simple device for comparing candle-powers which relies on the principle of equal illuminations. Two identical blocks of homogeneous paraffin wax are placed together as shown in Fig. 700 (i), with a sheet of tin foil or similar opaque material between them. This simple photometer is set up between the two

sources to be compared (Fig. 700 (ii)), with the tin foil at right angles to the line joining them. The light from each source enters the block of wax facing it and is diffused within the wax, causing the sides of the block to appear bright. The distances of the sources are adjusted until the eye, situated as shown in the figure, judges the sides of the two blocks to be equally bright. This condition obviously implies equal illumination of the two blocks, so that the ratio of the candle-powers is given by

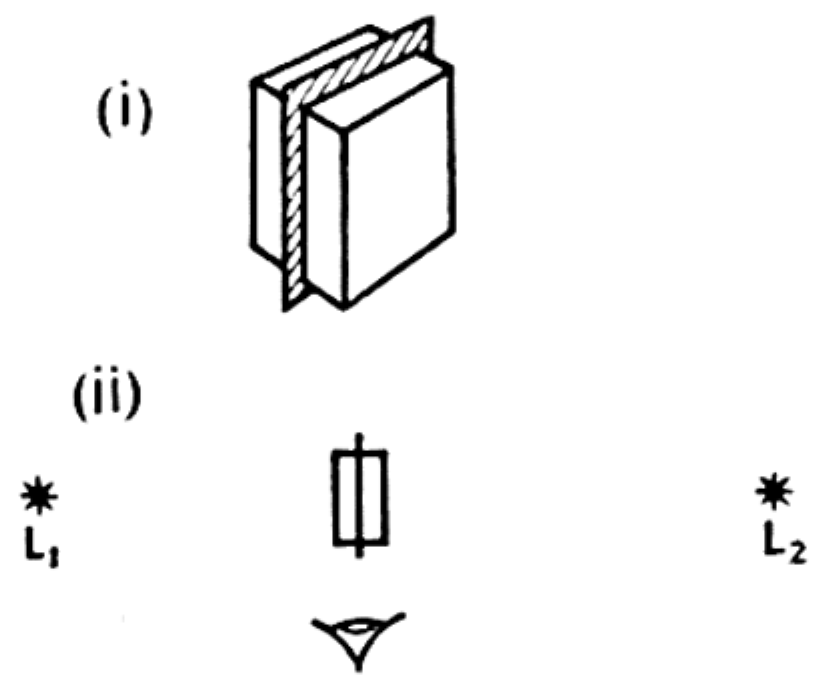


FIG. 700

$$\frac{I_1}{I_2} = \frac{d_1^2}{d_2^2}$$

where d_1 and d_2 are the respective distances of the sources from the photometer. The actual point on the photometer to which these distances should be measured is somewhat doubtful.

The Lummer-Brodhun Photometer.—This photometer is capable of considerable accuracy. It consists of a diffusing screen (A, Fig. 701 (i)) made of plaster of paris or some other suitable white substance, upon which the light from the two sources to be compared falls normally as indicated in the figure. Two mirrors M_1 and M_2 are placed so as to

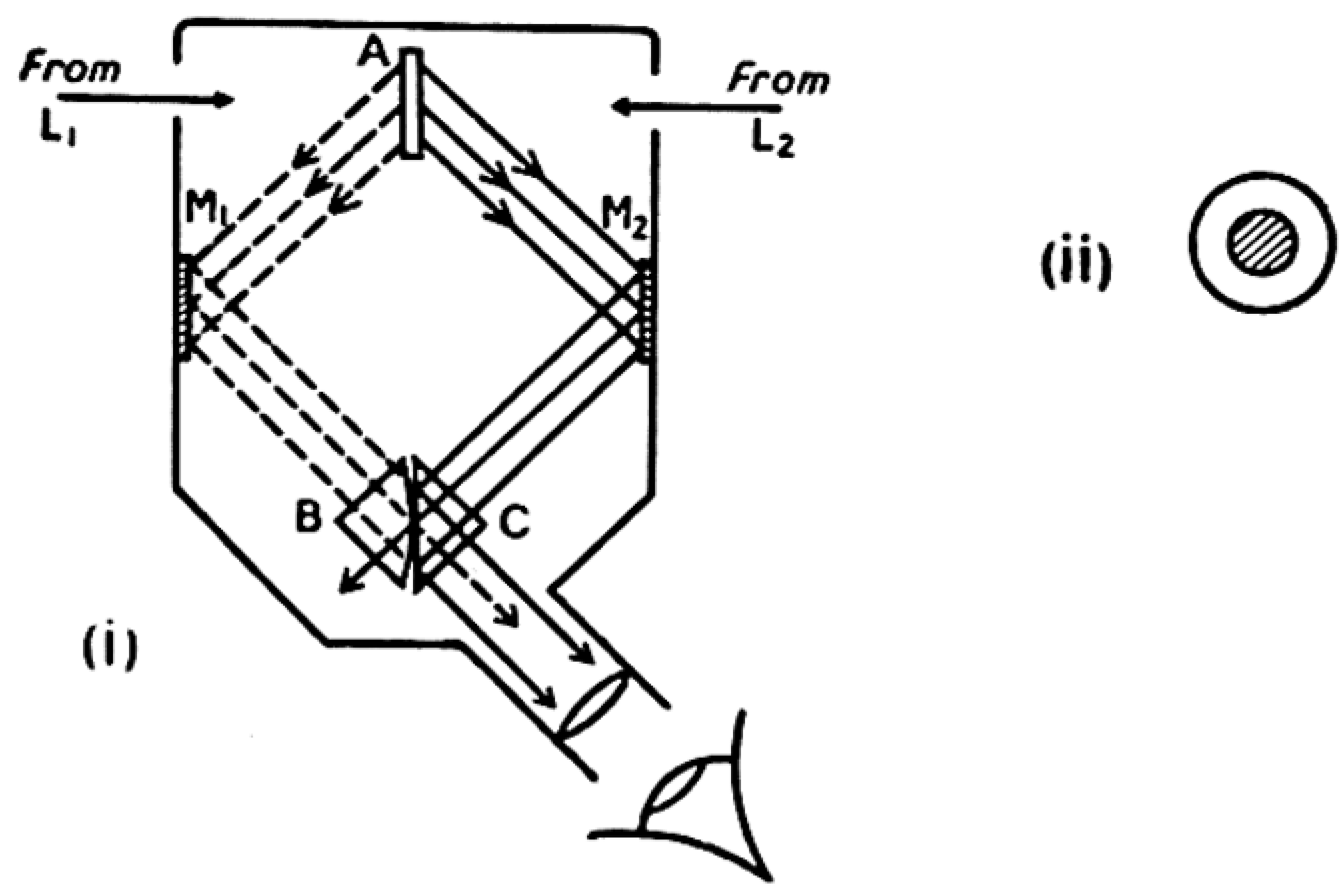


FIG. 701

reflect some of the diffused light on to the "cube," as it is called, which is a combination of two right-angled glass prisms. The hypotenuse face of one prism (B) is curved so that it is in contact with the corresponding face (which is flat) of the other prism (C) only in the central region. Therefore the light from the left-hand side of A entering B from M_1 is totally reflected except where the prisms are in contact. It passes on

through this region into C and thence into the eyepiece, which is the equivalent of a convex lens focused on the hypotenuse face of C. Thus light from the left-hand side of the diffusing screen constitutes the centre of the field of view. Light from the right-hand side of A makes up the surrounding region, because rays entering C from M_2 are totally reflected where the prisms are not in contact and so enter the eyepiece, while the central rays pass into B and are not seen. The appearance of the field of view is therefore like that shown in Fig. 701 (ii). The adjustment consists in moving the sources or the photometer until the two parts of the field are equally bright, so that the line of division between them is invisible. The distances of the sources from the screen A are then measured. To allow for the possibility that the two sides of the screen may have different reflecting factors the photometer is turned round (which has the same effect as interchanging the sources), the setting is repeated and the distances again measured. Equation (7) is then used to find the ratio of the candle-powers.

The Bunsen Grease-Spot Photometer.—This is a favourite instrument in elementary laboratories despite the fact that its theory is really more complicated than that of the Lummer-Brodhun type. It consists (Fig. 702) of a screen of thick white paper A having a matt (diffusing) surface. The screen is usually circular in shape and is stretched across a metal ring. A small spot of grease S is placed in the centre of the screen and renders the paper translucent here. Two plane mirrors are placed as shown in the drawing so as to enable both sides of the screen to be seen at once (Fig. 702 (ii)). The photometer is placed between the two sources to be compared, the screen being perpendicular to the line joining them.

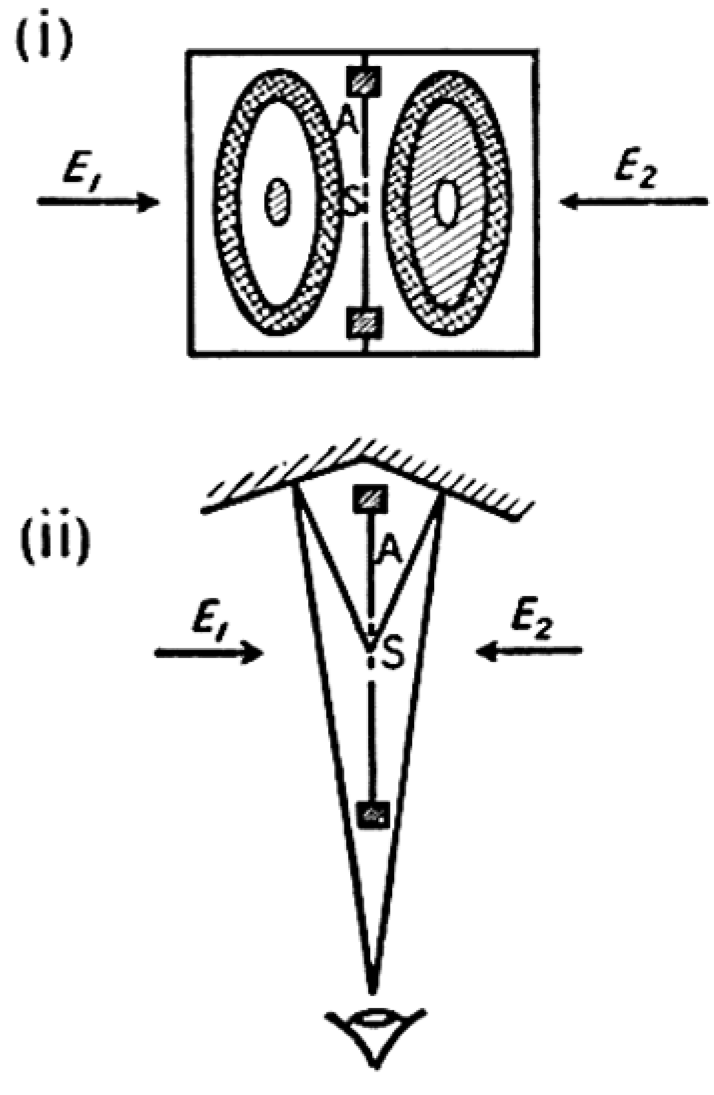


FIG. 702

Suppose that, for any given positions of the sources, the illumination of the left-hand side of the screen is E_1. The flux diffused by unit area of the opaque part of the screen will then be rE_1, where r is the reflecting factor for this part. Thus the brightness of this part of the screen will be proportional to rE_1 if we ignore the possibility of light coming through from the other side. For the grease spot there will be a corresponding quantity $r'E_1$, where r' will be considerably less than r. Suppose also that, owing to the illumination E_2 on the other side of the screen and the passage of light through the spot, a flux of $t'E_2$ emerges from the spot on the left-hand side. Then the brightness of the spot will be proportional

to $(r'E_1 + t'E_2)$. Similarly on the right-hand side, the brightness of the opaque part will be proportional to rE_2 and that of the spot to $(r'E_2 + t'E_1)$, supposing that the properties of the screen and · pot are the same on both sides.

Evidently the spot will appear as bright as the surrounding surface when viewed in the left-hand mirror if E_1 and E_2 have such values that

$$rE_1 = r'E_1 + t'E_2$$

If E_2 has a definite fixed value there is only one value of E_1 which will satisfy the above condition. It is therefore possible to compare two lamps by keeping a third lamp at a fixed distance on the right-hand side of the screen, thus keeping E_2 constant, and, in turn, placing each of the lamps to be compared on the left-hand side at such a distance that it causes the spot to disappear on the left-hand side. The value of E_1 is then the same in each case. Therefore the candle-powers of the lamps placed on the left-hand side are in the ratio of the squares of their distances from the screen when the spot disappears on the left-hand side.

Another method of using the photometer consists in placing the lamps which are to be compared on opposite sides of the screen and adjusting their distances until the spot disappears on, say, the left-hand side. We then have

$$rE_1 = r'E_1 + t'E_2 \quad . \quad . \quad . \quad . \quad (8)$$

Next, the distances are adjusted until the spot disappears when viewed from the right-hand side, and for this condition we can write

$$rE_2' = r'E_2' + t'E_1' \quad . \quad . \quad . \quad . \quad (9)$$

where E_1' and E_2' are the new values of the illumination on the left- and right-hand sides respectively.

Equation (8) can be written

$$\frac{E_1}{E_2} = \frac{t'}{r - r'} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

and equation (9) gives

$$\frac{E_2'}{E_1'} = \frac{t'}{r - r'} \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Combined together the two equations give

$$\frac{E_1}{E_2} = \frac{E_2'}{E_1'} \quad . \quad . \quad . \quad . \quad . \quad (12)$$

If the candle-powers of the two sources are I_1 and I_2, and their distances from the screen for disappearance of the spot on the left-hand side are d_1 and d_2 and for disappearance on the right-hand side are d_1' and d_2',

we have $E_1 = \frac{I_1}{d_1^2}$, $E_2 = \frac{I_2}{d_2^2}$, $E_1' = \frac{I_1}{d_1'^2}$, $E_2' = \frac{I_2}{d_2'^2}$, so that equation (12) gives

$$\frac{I_1}{I_2} \cdot \frac{d_2^2}{d_1^2} = \frac{I_2}{I_1} \cdot \frac{d_1'^2}{d_2'^2}$$

whence

$$\frac{I_1}{I_2} = \frac{d_1 d_1'}{d_2 d_2'}$$

Thus the ratio of the candle-powers can be found.

A third method by which the ratio of the candle-powers can be determined consists in adjusting the distances of the sources until the contrast in brightness between the spot and the surrounding screen is the same on both sides. It is in the making of this setting that the two plane mirrors are particularly useful. If we represent the contrast by the *ratio* of the brightness of the screen to that of the grease spot, it can be shown that the contrast is the same on both sides when the illumination is the same (*i.e.* when $E_1 = E_2$), so that the usual photometer equation is then applicable, namely

$$\frac{I_1}{I_2} = \frac{d_1^2}{d_2^2}$$

In actual practice a photometer setting can be made more precisely when it involves the judgment of equal contrast rather than of equal brightness. The equal contrast principle has been adopted in one form of Lummer-Brodhun photometer for this reason.

The Flicker Photometer.—A simple form of this instrument can be set up as shown in plan in Fig. 703. A is a white cardboard screen having a matt diffusing surface, and B is a disc of the same material with sectors cut out of it as shown in the small drawing. This is usually described as a Maltese cross, and it is mounted on a spindle so that it can be rotated at any chosen constant speed.

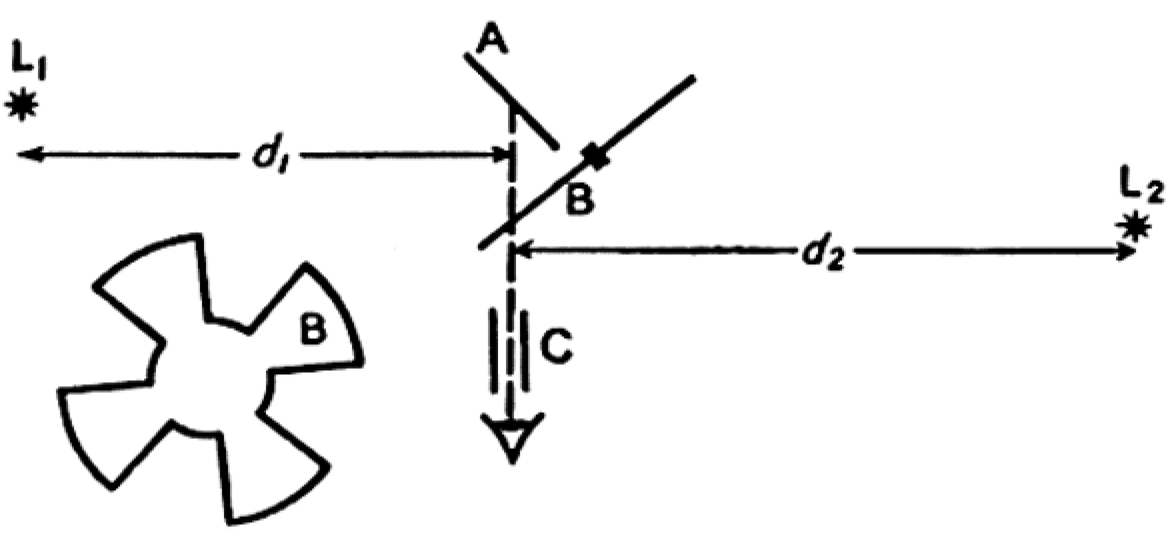


FIG. 703

The two sources to be compared, marked L_1 and L_2 in the figure (candle-powers I_1 and I_2), are arranged together with the screen A and the sectored disc B so that the light from L_1 falls on A at the same angle as that from L_2 falls on B. Both A and B are viewed through a tube or ow-power microscope C placed in such a position that the angle of viewing is the same in each case.

When B is rotating, the eye receives the diffused light alternately from

A and B. If these two surfaces are of different brightness a sensation of flicker is produced when B rotates slowly, but the flicker disappears when B is speeded up, because the eye cannot then register the alternations of brightness.

The flickering observed when B is run at a moderate speed can, however, be eliminated by adjustment of the distances of the sources from A and B. It is reasonable to suppose that when this adjustment has been made, the lack of flicker is due to the fact that A and B are equally bright, which means that they are receiving equal illumination. This can be verified by measuring the distances d_1 and d_2, when it is found that $\frac{d_1^2}{d_2^2}$ is equal to the value of $\frac{I_1}{I_2}$ found by some other form of photometer. In other words, the flicker photometer, which does not involve the direct judgment of equal brightness, gives the same result as a static photometer.

It has already been mentioned that flicker can be eliminated by rotating B sufficiently fast whatever may be the difference of brightness between A and B; that is to say, the arrangement is insensitive to contrast at high speeds of alternation. It is found that there is an optimum frequency at which the instrument is most sensitive, *i.e.* at which the difference of brightness which just causes flicker to appear is a minimum. This optimum frequency depends on the actual brightness. For normal working it is in the neighbourhood of 20 alternations per second. Also, the sensitivity of the flicker photometer increases with the brightness.

The flicker photometer owes its importance less to its use as an alternative to, say, the Lummer-Brodhun photometer than to its application to the comparison of sources of different colour. Such a comparison is almost impossible with a photometer which depends on the matching of brightnesses. When two sources of different colour are used with a flicker photometer and the disc B is gradually speeded up from rest, a stage is reached at which the eye is no longer conscious of the two separate colours of the surfaces A and B but, in general, a flicker is observed. This flicker can be eliminated by altering the distances of the sources. If they are d_1 and d_2 for absence of flicker, then the ratio of the intensities of the sources is taken to be given by

$$\frac{I_1}{I_2}=\frac{d_1^2}{d_2^2}$$

just as in the case of two sources of the same colour. This procedure can be justified to a fair degree of approximation by observations with the flicker photometer such as (i) when two lamps are placed close together so as to form the equivalent of a single source, the effective intensity is found to be the sum of the separate intensities; (ii) the ratio of the intensities of two lamps is found to be the same whether they are compared directly with each other or whether each is compared with a third lamp.

As with lights of the same colour, so with those of different colour, increasing the rate of alternation leads to a stage at which no flicker can be detected whatever may be the difference of brightness of the two surfaces.

The form of flicker photometer commonly found in teaching laboratories is that devised by Simmance and Abady. The essential feature of this instrument is a plaster of paris disc which can be regarded as being made up in the following way. Two truncated cones ABCD and EFGH (Fig. 704 (i)) are cut diagonally along AC and EG, and the parts ADC and EFG are stuck together on the faces AC and EG. Fig. 704 (ii) shows how this disc appears as it is rotated about the axis shown in the drawing. In the actual instrument this rotation is effected by a clockwork motor whose speed can be varied. The light from the two sources L_1 and L_2 falls near the edge of the disc in the directions shown in (*a*). At the stage depicted in (*a*) the eye receives only light originally coming from L_1. At (*c*), half a revolution later, it is only the light from L_2 which reaches the eye, while (*b*) and (*d*) are transitional stages.

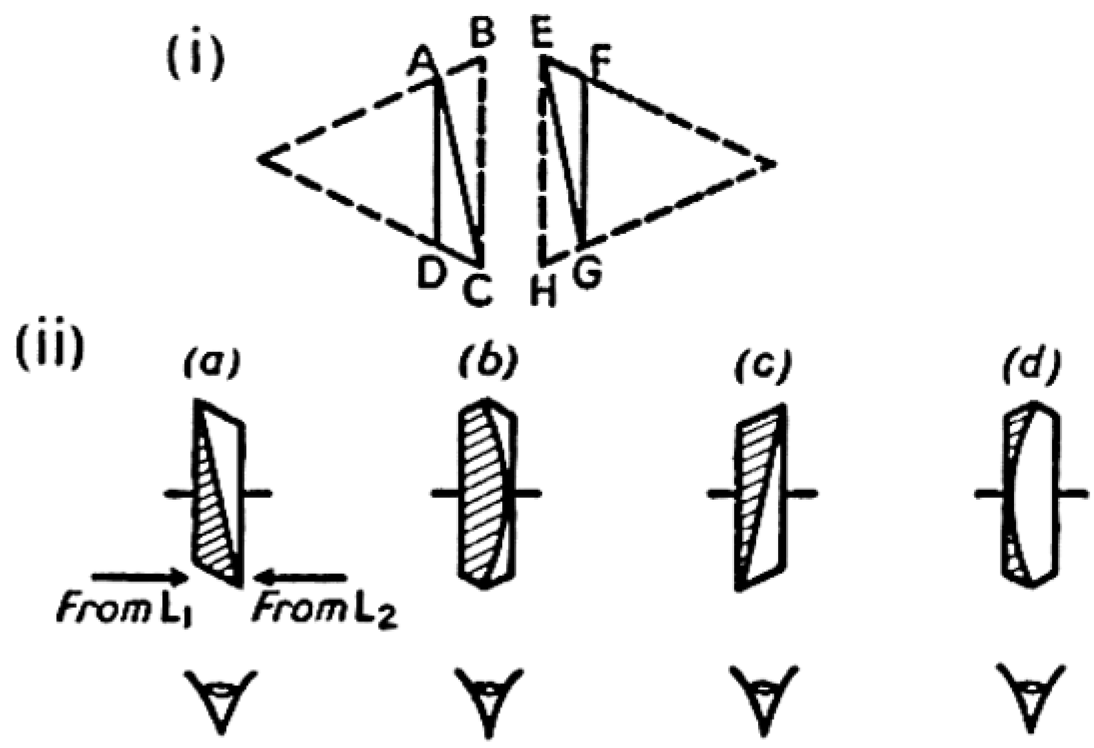


Fig. 704

Distribution of Light Flux round a Lamp.—It should be mentioned that the simple photometer equation (equation (4), page 906) is based on the assumption that the sources behave as point sources. However, it can be used with sufficient accuracy when the sources are, say, ordinary electric lamps, provided that they are at a distance of a metre or more from the photometer. Suppose that in a photometer experiment one of the sources is an electric lamp hanging in its usual position (*i.e.* with the cap at the top), and the other source is a lamp of known candle-power. After adjusting the photometer for equal brightness and measuring the distances, a result can be obtained by the usual equation for the candle-power of the unknown lamp. Now let this lamp be rotated about a vertical axis through successive angles of, say, 10°, and let its candle-power be determined for each setting. In general, the candle-power will be found to vary with the angle. What is determined in each case is the candle-power of the lamp in the direction of the line joining the lamp to the photometer. The results of such a series of determinations can be exhibited in the way shown in Fig. 705. The lamp may be supposed to be situated at O, and the diagram is constructed by plotting on each radial line a point such as P, whose distance from O is proportional to the

candle-power in the direction of the line. The curved line passing through all the points such as P shows the distribution of candle-power in the particular plane. The mean candle-power in this plane is the mean value of such lengths as OP for a large number of equally-spaced directions extending right round the complete 360°.

A full investigation of the test lamp would involve determining its candle-power in a large number of separate directions, not all confined to one plane. We can then imagine that the results of such an experiment are represented by a surface described round the lamp such that the candle-power in any particular direction is proportional to the distance in that direction from the lamp to the surface. If the candle-power in a given direction is I, then the flux emitted in a small solid angle $\Delta\omega$ in that direction is $I\Delta\omega$ (page 901). The total flux emitted will therefore be $\Sigma I\Delta\omega$, where the summation is taken over all the solid angles contained in the space surrounding the lamp. The **mean spherical candle-power** (I_0) of the lamp is defined as the candle-power of a point source emitting uniformly in all directions which would emit the same total flux as the lamp. This total flux is equal to $4\pi I_0$, so that I_0 is given by the equation

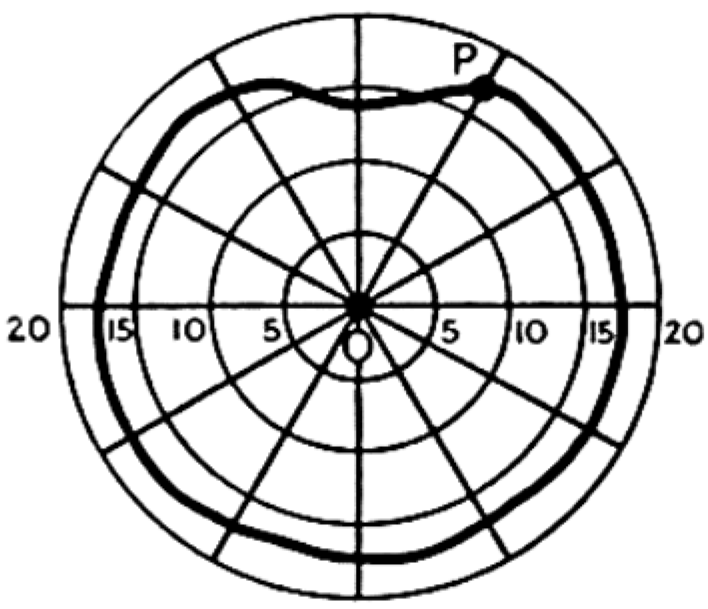


FIG. 705

$$4\pi I_0 = \Sigma I\Delta\omega$$

where $\Delta\omega$ tends to zero.

Suppose that the candle-power of a given lamp is determined in a series of n equally spaced directions. A solid angle of $\frac{4\pi}{n}$ will be associated with each of these directions, so that $\frac{4\pi}{n}$ can replace $\Delta\omega$ in the above equation, and, provided that n is very large (infinity to be exactly correct), we can find I_0 from the equation

$$I_0 = \frac{\sum \frac{4\pi I}{n}}{4\pi}$$

$$= \frac{\Sigma I}{n}$$

Thus I_0 is simply the arithmetic mean of the separate candle-powers, provided that the directions in which the candle-powers are determined are equally spaced and are very numerous.

The performance of such a series of separate determinations involves a great deal of time, and another method has therefore been devised which,

however, requires the use of a lamp for which the mean spherical candle-power has already been determined by the long method. It involves the use of what is called an **integrating sphere.** This is a large hollow sphere of 3 to 10 feet diameter, coated on the inside with a white matt diffusing surface. The lamp to be tested is placed inside the sphere. On account of the large number of diffuse reflections which the light leaving the lamp in any particular direction undergoes before it is all finally absorbed, the illumination of the surface is uniform and is proportional to the total flux from the lamp and therefore to the mean spherical candle-power.

A small opening is made in the wall of the sphere and a window W (Fig. 706) of some translucent material like ground glass or opal glass is placed in it with its inner surface flush with the inner surface of the sphere. The lamp L to be tested

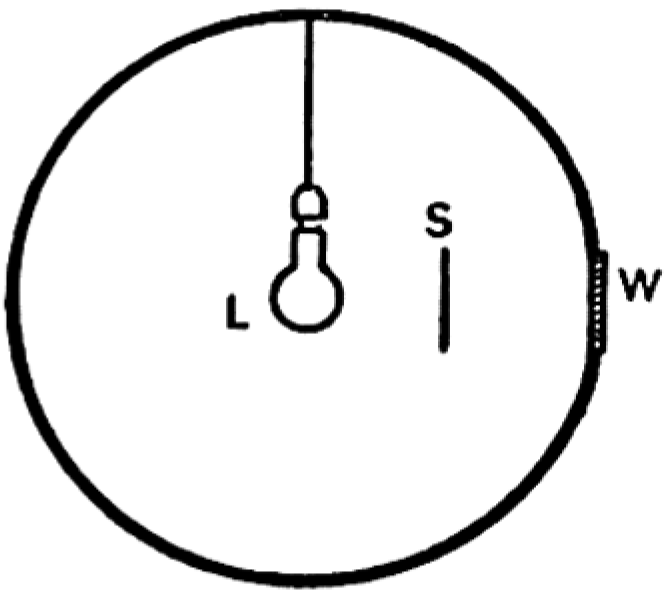


FIG. 706

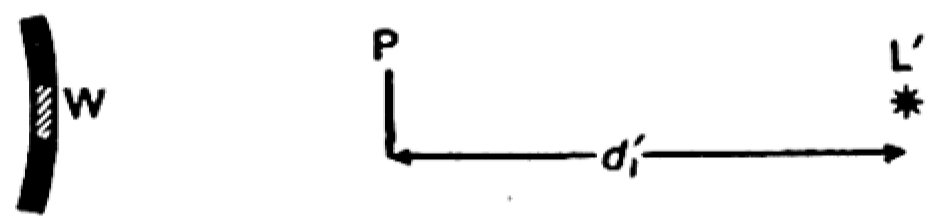


FIG. 707

is placed in the centre of the sphere and a small screen S prevents direct light reaching W from L.

The effective candle-power of the window, as viewed from outside, is compared with that of another lamp, L′, by means of a photometer. The arrangement is shown diagrammatically in Fig. 707, in which P represents the photometer. The distance of W from P is kept fixed and L′ is moved until there is equal illumination at the photometer. If the distance PL′ is then equal to d_1' and the candle-power of L′ is I', this illumination is equal to $\frac{I'}{d_1'^2}$. It is also equal to KI_0, where I_0 is the mean spherical candle-power of the lamp in the integrating sphere and K is a constant depending on the properties of the sphere and window and on the distance between W and P. Thus

$$KI_0 = \frac{I'}{d_1'^2}$$

The test lamp is then replaced by a lamp whose mean spherical candle-power I_s has been found from determinations of its candle-power in various directions. The lamp L′ is moved to a distance d_2' from P until there is again equal illumination at P. We then have

$$KI_s = \frac{I'}{{}_2d'^2}$$

so that, by division,

$$\frac{I_0}{I_s} = \frac{d_2'^2}{d_1'^2}$$

from which I_0 can be calculated.

Determination of Light Transmitted by a Plate.—Suppose that two sources, L_1 and L_2, of candle-powers I_1 and I_2 respectively, produce equal illuminations at a photometer P (Fig. 708 (i)) when their distances are d_1 and d_2 respectively.

Now let a plate, G, of glass or other transparent material be interposed between P and L_1, the distance d_1 being unaltered (Fig. 708 (ii)). Owing

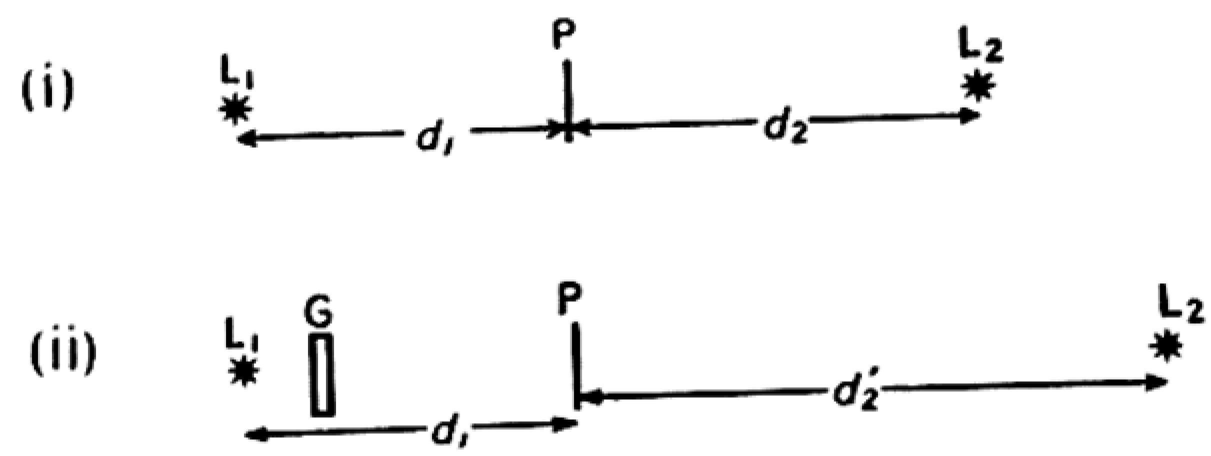


FIG. 708

to the reflection of light by the surfaces of the plate and to the absorption in passing through the material, the illumination on the left-hand side of P will be reduced, and to balance the photometer the source L_2 must be moved further from P, say to a distance d_2'. Therefore the illumination of the photometer screen (*i.e.* the flux per unit area) is $\dfrac{I_2}{d_2^2}$ without the plate and $\dfrac{I_2}{d_2'^2}$ with it. The fraction of the light flux which the plate transmits is therefore $\left(\dfrac{I_2}{d_2'^2} \div \dfrac{I_2}{d_2^2}\right)$ or $\dfrac{d_2^2}{d_2'^2}$, while the fraction cut off by the plate is $\left(1 - \dfrac{d_2^2}{d_2'^2}\right)$. It is, of course, equally possible to do the experiment by keeping d_2 constant and altering d_1.

Strictly speaking, the distance of L_1 from P is effectively reduced by the presence of the plate by $t\left(1 - \dfrac{1}{n}\right)$, where t is the thickness of the plate and n is the refractive index of the material (page 744). Allowance can be made for this either in the calculation or by increasing d_1 by this amount when the plate is present. The fraction of the light incident upon it which is actually absorbed (as distinct from reflected) by the material of the plate can be determined by doing two experiments of the kind described above with two separate plates of the same material having similar surfaces but different thicknesses. The fraction of the light which is absorbed by a thickness of material equal to the difference in thickness of the two plates is found by subtraction.

***Example.**—A horizontal table is illuminated by a 200 c.p. lamp held 6 ft. vertically above its centre. Calculate the intensity of illumination at the centre of the table. When a glass plate is placed below the lamp, the intensity of illumination is diminished by 19 per cent. and may be restored to its original value* (a) *by changing the candle-power of the lamp or* (b) *by changing its position. Calculate the new candle-power required in* (a) *and the new height above the table in* (b). (L.I.)

The illumination on a horizontal surface at a point 6 ft. vertically below a lamp of 200 candle-power is $\frac{200}{6^2}$ or 5·56 ft.-candles.

(*a*) Since the glass plate causes a diminution in the illumination of 19 per cent., the illumination at the point considered when the glass plate is present must be $\left(\frac{200}{36}\times\frac{81}{100}\right)$ ft.-candles. If a lamp of candle-power I is substituted for the original lamp so as to restore the illumination, then I must be given by

$$\frac{I}{36}\times\frac{81}{100}=\frac{200}{36}$$

from which we obtain

$$I=\frac{100\times 200}{81}\text{ candle-power}$$
$$=247\text{ candle-power approximately}$$

(*b*) If d is the distance at which the 200 candle-power lamp must be placed to give the same illumination as when the glass is absent and the lamp is 6 ft. above the table, we have

$$\frac{200}{d^2}\times\frac{81}{100}=\frac{200}{6^2}$$

from which we obtain

$$d=\sqrt{\frac{6^2\times 81}{100}}$$
$$=5\cdot 4\text{ ft.}$$

Determination of Light Reflected by a Surface.—This may be carried out as follows. Two sources L_1 and L_2 of candle-powers I_1 and I_2 respectively are arranged to balance a photometer P (Fig. 709 (i)), their

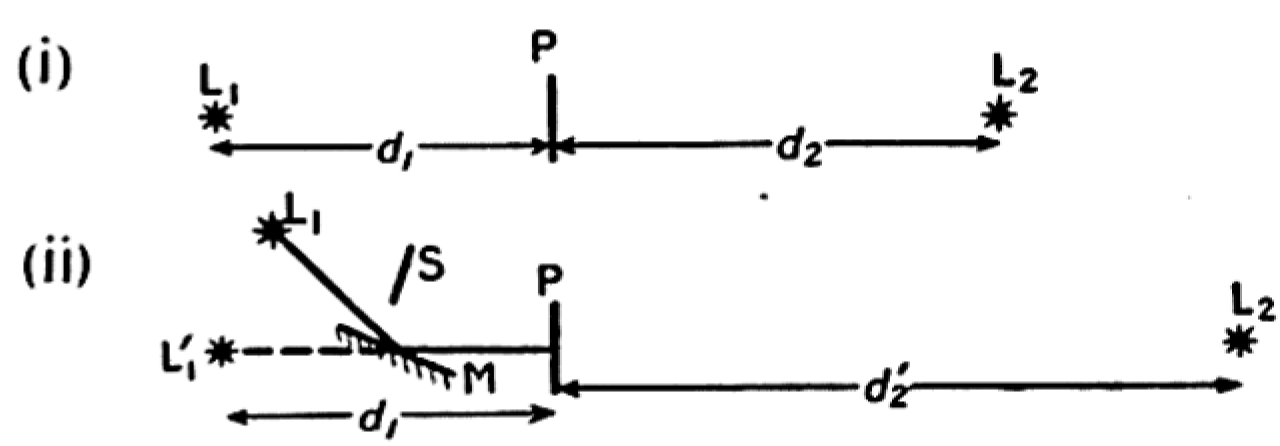


FIG. 709

distances from it being d_1 and d_2. The reflecting surface M is now placed between L_1 and P, and L_1 is moved to such a position that light from it is reflected by M and then strikes P in the same direction as before

(Fig. 709 (ii)). Also the distance of the new position of L_1 from M is made such that the length of the path of the light from L_1 to P is d_1 as before. This means that the image L_1' of L_1 is situated where the actual source was in Fig. 709 (i). A screen S must be placed so as to prevent light from L_1 reaching P directly. Owing to loss of light during the reflection, the illumination on the left-hand side of P is less than before, so that L_2 must be moved to a distance d_2' from P in order to balance the photometer. The illumination of P in (i) is $\frac{I_2}{d_2^2}$ and in (ii) it is $\frac{I_2}{d_2'^2}$. Therefore the fraction of the light falling upon it which M reflects is $\frac{d_2^2}{d_2'^2}$, and the fraction which is lost is $\left(1 - \frac{d_2^2}{d_2'^2}\right)$. Obviously the angle of incidence of the light on M can be altered, and the variation of its reflecting power with this angle can be investigated.

3. THE MEASUREMENT OF ILLUMINATION

The Macbeth Illumination Photometer.—It frequently happens that the illumination at any place in a building or street is due to more than one source of light and to reflections from walls, ceilings, etc. as well as to light received directly from the sources. It is therefore useful to have methods of measuring the illumination at any place without having to calculate it from a knowledge of the candle-power and positions of the sources and of the reflecting powers of the relevant surfaces.

One principle which is used for the measurement of illumination consists in placing a diffusing **test surface** at the place where the value of the illumination is required and, in effect, using a photometer to compare its brightness under these circumstances with its brightness when it is placed at a known distance from a source of known candle-power. The Macbeth illumination photometer (or illuminometer) is one of the instruments used for this purpose. It is shown diagrammatically in Fig. 710. The two prisms A, constituting the "cube" of a Lummer-Brodhun photometer, enable the eye to compare the brightness of the test surface with that of a piece of opal glass B illuminated by a lamp L situated at a variable distance behind it. The position of L is altered, thus changing the brightness of B, until a photometric balance is obtained, and the reading of the scale attached to the rod carrying L is then taken against a fixed mark. The instrument can be calibrated and the scale made to read in units of illumination of the test surface, such as foot-candles, by viewing the test surface when it is placed at a known distance d from a source of known candle-power I. Its illumination is then equal to $\frac{I}{d^2}$, and this is the figure to which the reading of the scale corresponds when

L has been adjusted to give a balance in this standardizing experiment. The distance between the test surface and the photometer does not affect the photometric balance, provided the test surface fills the field of view

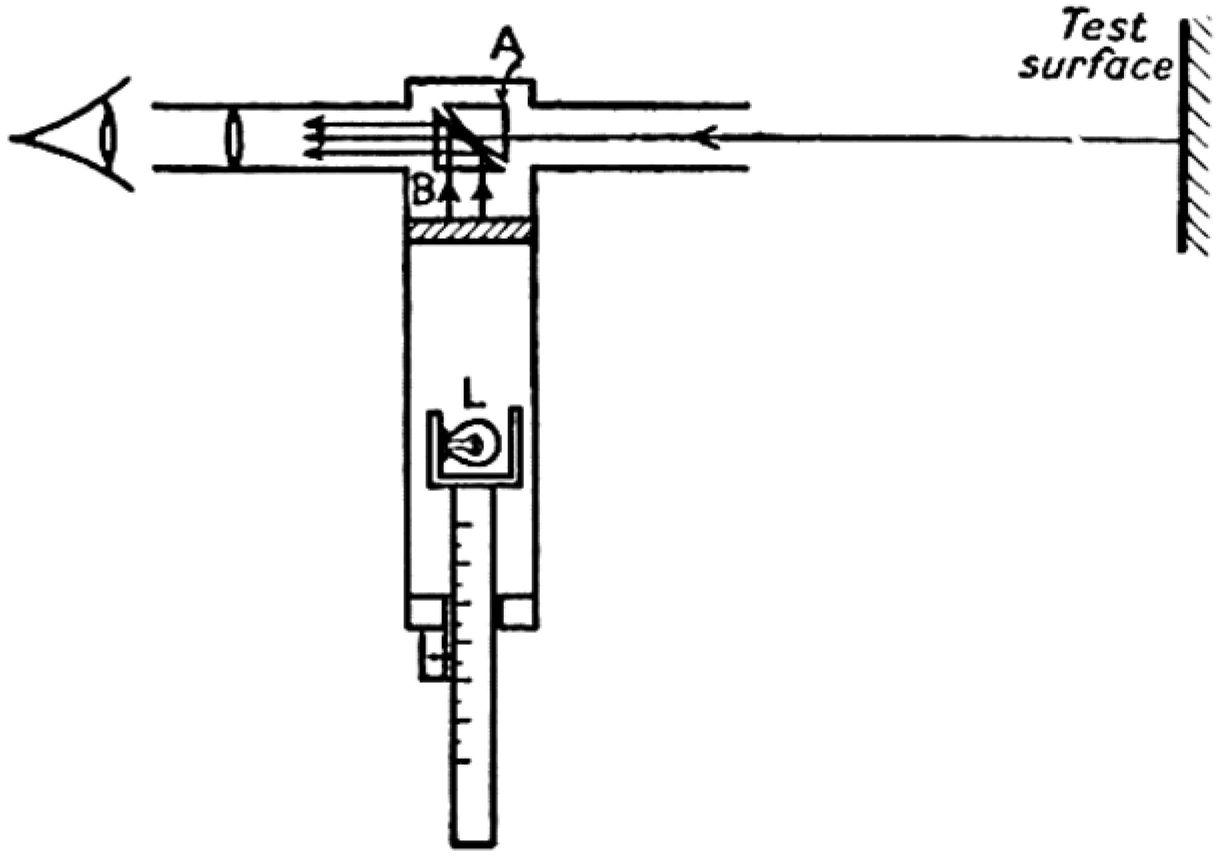


FIG. 710

of the photometer, because the brightness of a surface is independent of the distance from which it is viewed (page 905).

The Foot-Candle Meter.—This instrument can be used for rapid measurement but is not of very great accuracy. A strip of white paper (S, Fig. 711) carrying a row of Bunsen grease spots is set in the top of a box. At one end of the box is a small electric lamp which sends light through the spots both directly and by reflection at the strip of plane mirror M. The brightness of the spots viewed from outside the box will decrease progressively from the end nearest the lamp. When external light falls on the top of the box, one of the spots will be nearest in brightness to the surrounding opaque white paper, the brightness of which is due to the illumination which is being estimated. The illumination can then be found by reading the position of the least visible spot against a scale of foot-candles alongside the row. The scale is obtained by calibration. The instrument incorporates a dry battery for lighting the lamp, a voltmeter and rheostat, and before use the voltmeter must be set to the reading at which the foot-candle meter was calibrated. Small readjustments of the instrument can be made by altering the inclination of the mirror M.

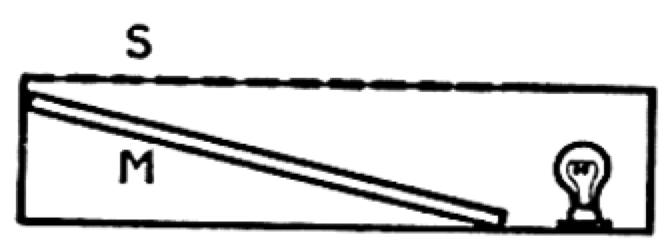


FIG. 711

The Photo-Cell.—Much illumination measurement is done nowadays by means of instruments incorporating photo-electric cells, or photo-cells as they are often called. The photo-electric effect and the construction of photo-cells is described in Vol. 5 of this book. It is sufficient now to explain that light falling on the surface of a metal causes electrons to

be emitted which constitute a measurable electric current if they are collected by an electrode. An instrument for measuring illumination based on this principle therefore contains the photo-cell and a galvanometer which gives a reading proportional to the photo-electric current. The instrument can evidently be calibrated with the help of a lamp of known candle-power. One of the commonest forms of the instrument is the **exposure-meter** used in estimating the illumination of a scene or object during photography.

Illumination Engineering.—This is the name given to the applied science which deals with all aspects of illumination. One of the very important matters with which it is concerned is the correct degree and distribution of illumination in rooms, offices, factories, halls, streets, etc. For general illumination a value of, say, between 5 and 20 foot-candles (lumens $ft.^{-2}$) is usually aimed at according to the activity to be carried on in the room or space concerned. Ordinary reading can be done in comfort with 5 foot-candles, while close work may require 25 foot-candles, and even up to 100 or more for a surgical operation.

The **efficiency** of a source may be defined as the quotient of the flux in lumens which the lamp gives by the (electrical) power in watts which it draws from the mains supply. It can also be expressed as candle-power per watt, which will be $1/4\pi$ of its value in lumens per watt. It is also a common practice to express efficiency the other way round, *i.e.* as watts per lumen or watts per candle-power. The lumens given per watt by a tungsten filament gas-filled lamp increases very much with the size (wattage) of the lamp. For small lamps the proportion of power dissipated as heat by conduction and convection is greater. This loss of heat is unavoidable, and constitutes a very large wastage of power in lamps of the incandescent filament type because of the large amount of non-visible (infra-red) radiation which is necessarily emitted. By operating the lamp at the highest possible filament temperature the proportion of visible radiation to infra red is increased and the lamp is more economical (page 529, Vol. 2).

Lamps in which an electric discharge passes through a rarefied gas are more economical than filament lamps, especially if the non-visible (ultra-violet) radiation which may be emitted from the discharge is converted into visible light by coating the tube with a material which fluoresces, *i.e.* gives out visible light when exposed to ultra-violet radiation. Modern "strip" lighting uses this principle.

EXAMPLES L

1. Explain how to compare the candle-powers of two point sources of light.

If the total light flux from a point source is 1232 lumens, calculate from first principles (*a*) the candle-power of the source, (*b*) the illumination on a small surface which is 4 ft. from the source and inclined at an angle of 30° to the light rays incident upon it. (L.I.)

2. Describe how the candle-powers of two lamps may be compared.

A horizontal sheet of notepaper is illuminated by a 50-c.p. lamp 2 ft. above it and a 200-c.p. lamp 6 ft. vertically above the first lamp. Calculate the candle-power of the lamp which must replace the upper one if the lower one is dispensed with and the intensity of illumination of the paper is to be unaltered. (L.I.)

3. Explain the terms *lumen*, *foot-candle*.

Describe and explain a method for comparing the intensities of two sources of somewhat different colours.

Two lamps each of 1000 candle-power are suspended 20 ft. above the ground at a distance apart of 96 ft. Find the illuminations at a point on the ground (*a*) directly beneath one lamp, (*b*) midway between the lamps. (L.I.)

4. Define *illumination* at a point on a surface. In what units is it commonly expressed?

A small source of 100 candle-power is suspended 3 ft. vertically above a point P on a horizontal surface. Calculate the illumination at a point Q on the surface 4 ft. from P.

Either (*a*) prove any formula you employ in your calculation, *or* (*b*) describe an experimental method for finding the variation of illumination over the surface. (L.I.)

5. Derive an expression connecting the illumination at a point on a surface with the distance of the point from a small source of light and the angle which the line joining source and point makes with the normal to the surface.

Describe a method for finding the illumination on a horizontal surface at different points in a room. (L.I.)

6. Define illumination (intensity of illumination) and luminous intensity (illuminating power).

Describe a good method of comparing the illuminating powers of two lamps of different colour. What is the chief disadvantage of the method?

A lamp of 100 candle-power hangs 5 ft. above a table. On the ceiling, 1 ft. above the lamp, is a plane mirror which reflects 50 per cent. of the light falling upon it. Calculate the intensity of illumination (*a*) at a point A on the table immediately below the lamp, (*b*) at a point B which is 4 ft. to one side of A. (C.H.S.)

7. In order that the relay of a burglar alarm can be actuated by means of a certain vacuum photo-cell, the latter must receive a minimum luminous flux of 0·05 lumen. If the cell is shrouded except for a rectangular aperture 0·3 cm. by 3·0 cm., determine the greatest distance at which a point source of 50 candles can be situated for the relay to operate. Calculate the distance through which the lamp must be moved if a glass plate of light transmission coefficient 0·64 is interposed between the lamp and the cell. (L.A., abridged.)

8. Describe an accurate form of photometer for comparing the candle-powers of lamps.

A lamp is 100 cm. from one side of a photometer and produces the same illumination as a second lamp placed at 120 cm. on the opposite side. When a lightly smoked glass plate is placed before the weaker lamp, the brighter one has to be moved 50 cm. to restore the equality of illumination. Find what fraction of the incident light is transmitted by the plate. (L.I.)

9. Describe and explain how you would proceed to:

(*a*) compare the candle-powers of two lamps;
(*b*) determine the ratio of the luminous intensities of the emergent and incident beams when light is passed through a glass plate;
(*c*) arrange two lamps of unknown and unequal candle-powers so as to produce on a screen two areas of illumination such that the intensity of illumination in one area is double that in the other. (L.I.)

10. Define *lumen*, *foot-candle*, and show how they are related.
Describe and explain how you would make an accurate comparison of the illuminating powers of two lamps of the same type.
A photometric balance is obtained between two lamps A and B when B is 100 cm. from the photometer. When a block of glass G is placed between A and the photometer, balance is restored by moving B through 5 cm. Where must B be placed in order to maintain the balance when *two more* blocks, identical with G, are similarly placed between A and the photometer? (L.H.S.)

11. Describe some form of photometer and its use.
A lamp is situated 10 cm. in front of a plane mirror, and a screen must be placed 100 cm. from the mirror in order that the lamp and its image may produce the same intensity of illumination at the screen as a lamp equal in every respect to the first lamp but situated 70 cm. from the screen. What percentage of the light falling upon it is reflected by the mirror? (L.I.)

12. Describe one form of photometer and its use, mentioning the precautions needed to ensure accuracy.
A 16-candle-power lamp is placed at the centre of a hollow sphere 4 ft. in diameter. What will be the intensity of illumination of the inner surface of the sphere (*a*) if this is dead black, (*b*) if it is white and reflects diffusely 80 per cent. of the light falling on it? (L.I.)

13. What is meant by the illumination of a surface? How would you show experimentally that the illumination of a surface receiving light normally from a small source varies inversely as the square of the distance from the source?
A parallel beam of light is obtained by placing a small source of light of 20 c.p. at the principal focus of a convex lens of focal length 20 cm. and diameter 10 cm. The beam falls obliquely on a screen, at an angle of incidence of 30°. Find, in metre-candles, the illumination of the screen, if the lens transmits 80 per cent. of the light falling on it. (O.H.S.)

14. What is meant by illuminating power and intensity of illumination? How are they related?
A small source of 32 c.p. giving out light equally in all directions is situated at the centre of a sphere of 8 ft. diameter, the inner surface of which is painted dead black. What is the illumination of the surface?
If the inner surface is repainted with a matt white paint which causes it to reflect diffusely 80 per cent. of all light falling on it, what will the illumination be? (L.H.S.)

Chapter LI

INTERFERENCE OF LIGHT

1. THE PRODUCTION AND OBSERVATION OF THE INTERFERENCE OF LIGHT

In Section 1 of Chapter XXXVI (Vol. 3) there is a discussion of the interference of waves, mainly in connection with sound. It is explained there that when two sources of sound are emitting notes of the same frequency there are places where the two waves are always out of phase with each other so that each tends to cancel the other, and the net effect is a weaker sound than anywhere else in the region. On the other hand, the sound is a maximum at places where the two waves are arriving in phase with each other.

If light is a wave motion we should expect it to exhibit similar interference phenomena. It must be said immediately, however, that interference is not observable when two *independent* light sources are used. This is due to the fact that the sum total of the light emitted by, say, an incandescent metal or vapour (as in a sodium flame) is made up of the small amounts of light given out independently by each separate atom of the substance and, moreover, each atom emits discontinuously, so that the phase of the composite wave due to all the separate atoms is constantly suffering abrupt changes. Therefore there is no definite and constant phase difference between the waves arriving at any one point from the separate sources, and sustained interference cannot take place. However, interference can be obtained between a light source and its image or between two images of the same source, because the two waves will then undergo the same phase changes simultaneously, and the phase difference between them at any one point will not vary with time. Light sources which will give rise to interference are said to be **coherent sources.**

Fresnel's Biprism.—This is a simple arrangement for the production of two images of the same source. Its principle is illustrated in Fig. 712. The biprism can be regarded as a glass prism having a refracting angle at A of nearly 180°. Its name comes from the fact that it functions as two very thin prisms placed base to base. If a point source S is placed on the bisector of the angle at A, the bent line SAB_1 is the path of the ray from S which strikes the biprism just above A and so is bent downwards forming a virtual image of S at S_1, while SAB_2 is the corresponding ray for the lower half and gives the image S_2. Every point in the shaded region B_1AB_2 receives light from both halves of the prism and can be regarded

as receiving light direct from two coherent sources S_1 and S_2. In order to observe the resulting interference effects in practice the source S must cover only a small space in the plane of the paper in Fig. 712, otherwise any one point in the shaded region will be at different distances from the various parts of S_1 and S_2. We can, however, extend S (and therefore S_1 and S_2) in a direction perpendicular to the paper, which we do by using as the source S a slit in a screen with, say, a sodium flame behind it. The slit must be parallel to the edge A of the biprism, and S_1 and S_2 are then virtual images of the slit.

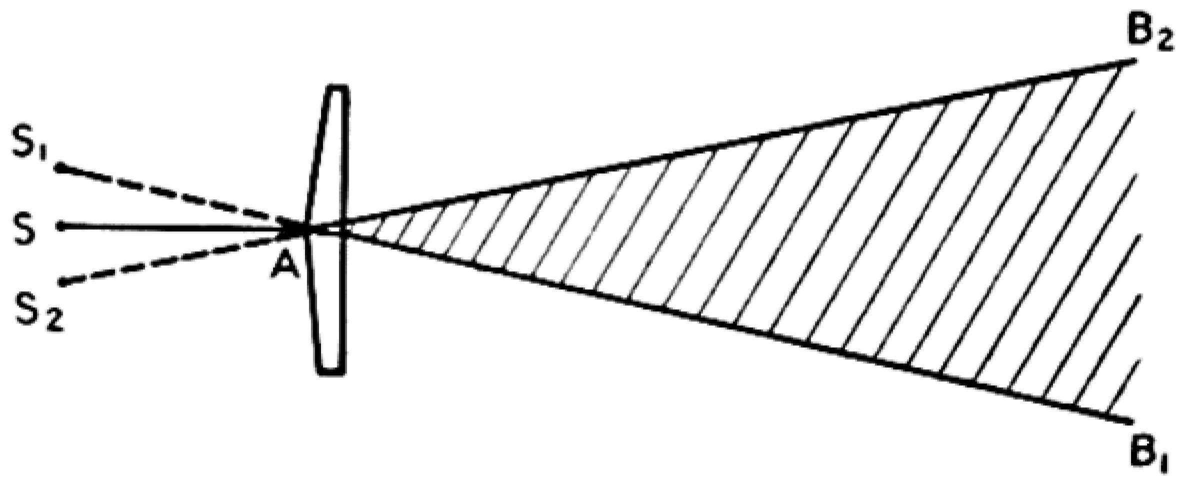


FIG. 712

We now examine the type of interference pattern obtained with the above arrangement. Let P (Fig. 713) be any point in the shaded region in Fig. 712. The difference in the lengths of the paths by which the light from S_1 and S_2 would reach P if they were actual sources can be found by drawing the arc S_1Q with centre P so that $QP = S_1P$. The path difference is then equal to the short length S_2Q. The point B is midway between S_1 and S_2, BC is the bisector of the angle A of the biprism and PC is perpendicular to this line. The distance PC is denoted by x and S_1S_2 by d. When the experiment is set up, the plane in which the interference is observed, *i.e.* PC, is at a considerable distance, say a metre, from S_1 and S_2, and the distances d and x are small in comparison, being only a few millimetres. Therefore the arc S_1Q can be regarded as a straight line which is perpendicular to all three lines S_1P, BP and S_2P, and hence we can consider S_1S_2Q and BPC as being similar triangles, so that

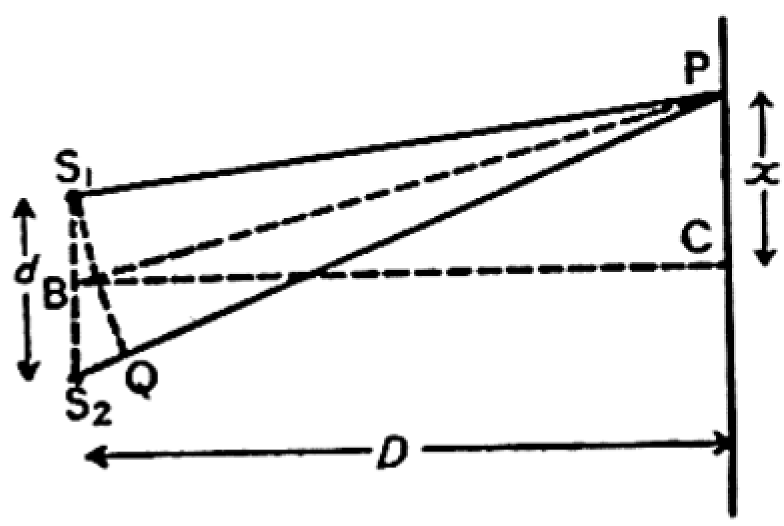


FIG. 713

$$S_2Q = \frac{xd}{BP}$$

$$= \frac{xd}{D}$$

where D stands for BP and can be taken in practice to be the distance between the slit and the plane where the interference is observed, since B is not very far removed from the slit S.

If a screen is held at a distance D from the slit there will be maximum illumination at a point whose position on the screen causes S_2Q, the path

difference, to be a whole number of wave-lengths of the light used. Therefore the values of x for points at which there is maximum light on the screen are given by

$$S_2Q = m\lambda$$

i.e.

$$\frac{xd}{D} = m\lambda$$

or

$$x = \frac{m\lambda D}{d}$$

where m can be 0 (for the point C), 1, 2, 3 . . . When we remember that the rays appearing to come from S_1 and S_2 are not confined to one plane and also that S_1 and S_2 are actually images of a *slit*, we realize that there will be *straight lines* of maximum illumination on the screen, known as bright **fringes,** whose directions are all parallel to the illuminated slit. Consecutive bright fringes are separated by dark fringes for which the path difference is an odd number of half wave-lengths. Thus the distance x from C of each dark fringe is given by the equation

$$x = \frac{(m + \frac{1}{2})\lambda D}{d}$$

where m is 0, 1, 2, 3 . . .

It is evident that the distance between any two consecutive bright (or dark) fringes is equal to the increase of x due to an increase of unity in the value of m, that is to say $\frac{\lambda D}{d}$. Thus the fringes are equally spaced, because this expression is independent of the position of the fringe, at any rate to a first approximation.

So far, we have imagined the fringes to be formed on a screen, but for the purpose of measuring their separation and so calculating λ it is usual to observe them by means of a microscope or simply an eyepiece. The simplest form of the latter is shown in Fig. 714. It consists of a converging lens L (or better, an eyepiece of the Ramsden type) mounted at one end of a tube, at the other end of which are crosswires W. For the sake of robustness these are often two perpendicular scratches on a sheet of plane glass. To adjust the eyepiece the lens is moved to and from the crosswires until they are in focus for an eye situated as shown in the drawing. If the eyepiece is then placed with its axis on or near the bisector of the angle of the biprism and directed towards the latter when the slit S is also on this line, the fringes will be seen. When the eye is looking at the crosswires, the fringes seen are situated in the plane of the latter. There is no question of focusing the fringes because they

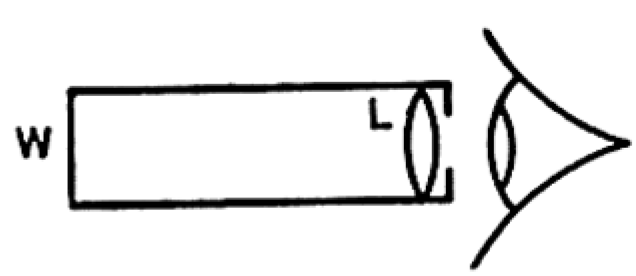


FIG. 714

exist everywhere in the shaded region in Fig. 712, and not in any particular plane. For this reason the fringes are said to be **non-localized.**

In order to measure the separation of the fringes, the eyepiece is mounted in such a way that it can be moved like a travelling microscope across the fringes by means of a micrometer screw. Readings of the screw are taken when the crosswires are on selected fringes, and the mean separation of, say, the bright fringes can be calculated from the readings. Since the separation of two consecutive fringes is equal to $\frac{\lambda D}{d}$ it is necessary to measure D, the distance of the crosswires from the slit, it being assumed that S, S_1 and S_2 are all in the same plane. It is also necessary to know d. the distance between the two slit images S_1 and S_2. This may be found in several ways. If a fairly powerful converging lens is interposed between the eyepiece and the biprism, its position can be adjusted until clear images of the two slits are seen in the plane of the crosswires. These are real images formed in the usual way, and in general there are two positions of the lens for which such images are formed (page 837). In one position the separation of the real images (d_1), which can be measured by the eyepiece, is less than S_1S_2 (*i.e.* d), and for the other position the separation (d_2) is greater than d. With the help of the analysis on page 837 the student should be able to show that $d=\sqrt{d_1d_2}$, from which d can be found. Thus λ can be calculated. The biprism experiment is *not* one of the most accurate ways of determining the wave-length of light.

If white, instead of monochromatic, light is used, all possible wave-lengths within the visible region are present and, in effect, the fringe systems due to all of them will be superimposed. These systems will not coincide with each other except at the central fringe, because the fringe separation depends on wave-length. Thus there is one central white fringe where the path difference is zero so that there is a bright fringe for all wave-lengths, but the fringes on each side of this are not black but red because they represent the minima for violet light, the absence of which from white light leaves a red coloration. Beyond the first few fringes, which are variously coloured for similar reasons, the illumination becomes increasingly uniform on account of the complicated overlappings of the fringes due to the various colours.

Other methods of producing the same type of interference as with a biprism will now be mentioned briefly, although these are not so common as the biprism in teaching laboratories.

Lloyd's Mirror.—This experiment is illustrated in Fig. 715. A slit S is illuminated by monochromatic light. A plane glass surface is arranged parallel to the slit in such a position that light from the slit strikes the surface at almost grazing incidence. The image of the slit is therefore situated at S′, quite close to S, and these two act as coherent sources to produce interference on, say, a screen situated at PC. The interference at a point such as P takes place between the direct ray SP and the ray SAP

which is reflected at A. The path difference can be calculated in the same way as for the biprism by regarding S and S′ as two coherent sources. The distance between consecutive fringes is given by the same formula as for the biprism, but there is an important difference in the fringe system due to the fact that when a light wave is reflected at the boundary of a medium of higher refractive index than the one in which it is travelling, it suffers a change of phase of π. This phenomenon is akin to the same phase change which a sound wave undergoes when it is reflected at a fixed boundary (page 593, Vol. 3). The change does not occur when the incident wave is travelling in the medium of higher refractive index.

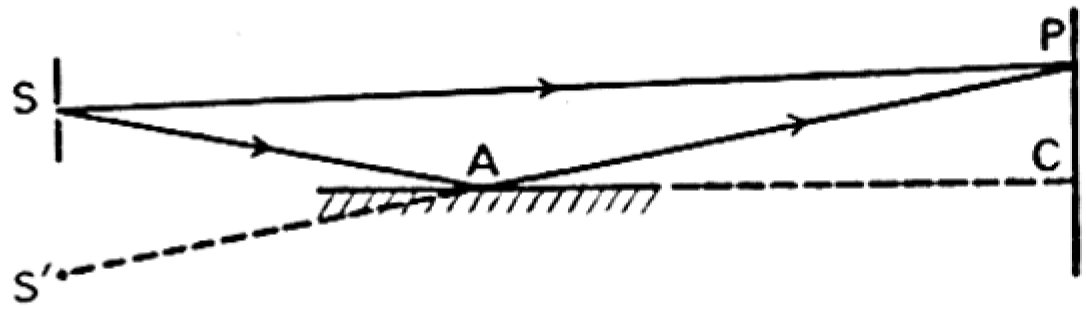


FIG. 715

The effect of the phase change is, of course, that when the position of a point such as P (Fig. 715) makes the actual difference of the distances SP and (SA + AP) equal to a whole number of wave-lengths, the two waves arrive at P in *opposite* phase and there is darkness. Similarly there are bright fringes where the difference in the geometrical path is an odd number of half wave-lengths. Thus the fringe at the point C, where the geometrical path difference is zero, is dark, whereas it is bright in the case of the biprism.

Fresnel's Mirrors.—In this experiment light from a slit is reflected separately by two plane reflecting surfaces placed near together and almost parallel to each other. The arrangement can be regarded as an extension of Lloyd's single mirror. Interference takes place between the two reflected waves, that is to say, the two images behave as coherent sources.

Young's Slits.—This experiment relies on diffraction, and indeed some of the effects of diffraction are seen in addition to those due to interference. A single slit S with a monochromatic source behind it is placed behind two parallel slits S_1 and S_2 situated close together in another screen as shown in Fig. 716. The light reaching S_1 and S_2 from S causes them to act as secondary sources so that each sends light to the screen. The two secondary sources are coherent because they are both derived from the single source S, and therefore interference will occur on the screen, giving light and dark fringes parallel to the slits. There will be a light fringe at P if the difference of the paths SS_1P and SS_2P is a whole number of wave-lengths, and a dark fringe if the difference is an odd number of half wave-lengths.

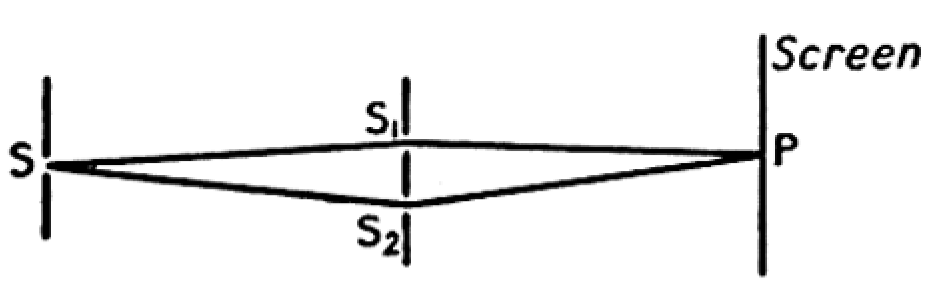


FIG. 716

2. INTERFERENCE DUE TO PLATES AND THIN FILMS

The General Principle.—Suppose that, as shown in Fig. 717, a ray of monochromatic light travelling in air is incident at A at an angle i on a parallel-sided plate of material of refractive index n. At A some of the incident light is reflected along AD, while some is refracted into the plate and strikes the opposite face at B. Refraction and partial reflection occur at both B and C as indicated in the figure. Owing to its having passed twice across the plate, the light travelling along CE has traversed a longer path than that along AD, but in calculating the path difference it is necessary to take into account the different refractive indices of the two media through which the separate rays have passed.

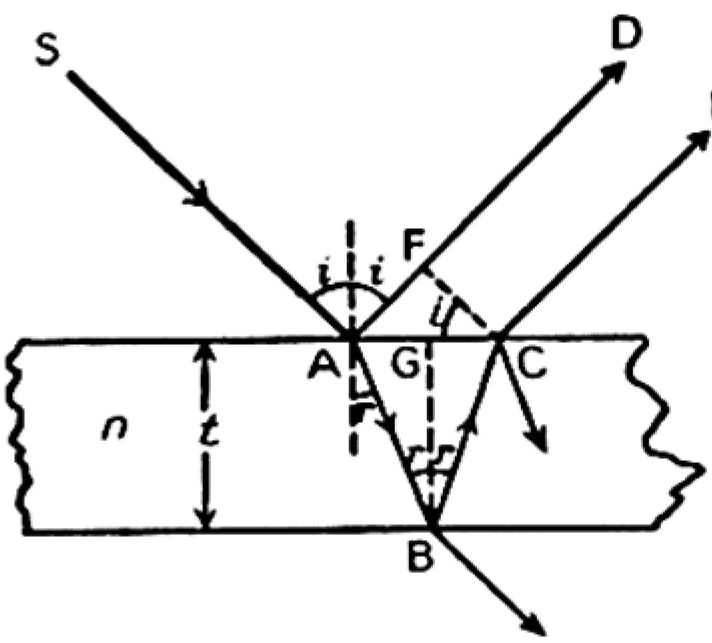


FIG. 717

The line CF is drawn perpendicular to the two rays into which the original incident ray has been separated. From the points C and F respectively the rays travel forward without any further relative phase change, so that the phase difference which exists between them is due to one ray having travelled from A to F through the air while the other went by the path ABC through the plate. If c and c_n respectively are the speeds of the light in air and in the plate, and if λ and λ_n respectively are the wave-lengths in air and in the plate, then, since the frequency of the light waves is the same in both media, we have

$$\frac{c}{\lambda}=\frac{c_n}{\lambda_n}$$

so that

$$\frac{\lambda_n}{\lambda}=\frac{c_n}{c}=\frac{1}{n}$$

as explained on page 696. Therefore

$$\lambda_n=\frac{\lambda}{n}$$

Also, if i is the angle of incidence at A and r is the angle of refraction,

$$\sin i=n\sin r$$

Furthermore, it is easy to see that $\widehat{ACF}=i$, and if BG is normal to the surfaces of the plate, $\widehat{ABG}=\widehat{CBG}=r$. Let GB, the thickness of the plate, be t. Then

$$AB=BC=\frac{t}{\cos r}$$

so that

$$AB + BC = \frac{2t}{\cos r}$$

The number of waves in the path ABC is equal to $\left(\frac{AB + BC}{\lambda_n}\right)$, *i.e.*

$$\frac{2t}{\lambda_n \cos r}$$

or

$$\frac{2nt}{\lambda \cos r}$$

Turning to the path of the light which is reflected back into the air at A we can write

$$\begin{aligned} AF &= AC.\sin i \\ &= 2t.\tan r.\sin i \\ &= 2nt.\frac{\sin r}{\cos r}.\sin r \\ &= 2nt.\frac{\sin^2 r}{\cos r} \end{aligned}$$

and the number of waves in the distance AF is equal to

$$\frac{2nt \sin^2 r}{\lambda \cos r}$$

Hence, apart from any phase change which may occur on reflection, the difference in the number of waves between the paths ABC and AF is

$$\frac{2nt}{\lambda \cos r} - \frac{2nt \sin^2 r}{\lambda \cos r}$$

which is equal to

$$\frac{2nt}{\lambda \cos r}(1 - \sin^2 r)$$

or

$$\frac{2nt}{\lambda}.\cos r$$

This is equivalent to saying that the retardation of the waves which go by the path ABC relative to those leaving A at the same time and going directly along AD is the same as the retardation due to a difference of path *in air* of $2nt \cos r$.

In actual fact, as has already been mentioned, a phase change of π occurs on reflection at A but not at the other reflections (or refractions) shown in Fig. 717. In terms of path difference this phase change is

represented by $\frac{\lambda}{2}$. Therefore the path difference (in air) between the two rays AD and CE must be written as

$$(2nt \cos r) - \frac{\lambda}{2}$$

Parallel Plates.—The path difference produced by a parallel plate or by a film of air bounded by two plane parallel sheets of glass (in which case $n = 1$) can be made use of to produce interference effects. Two rays cannot interfere unless they are made to meet each other. This can be brought about by inserting a converging lens in their path as indicated in Fig. 718, in which case they are brought to a focus at P in the focal plane of the lens.

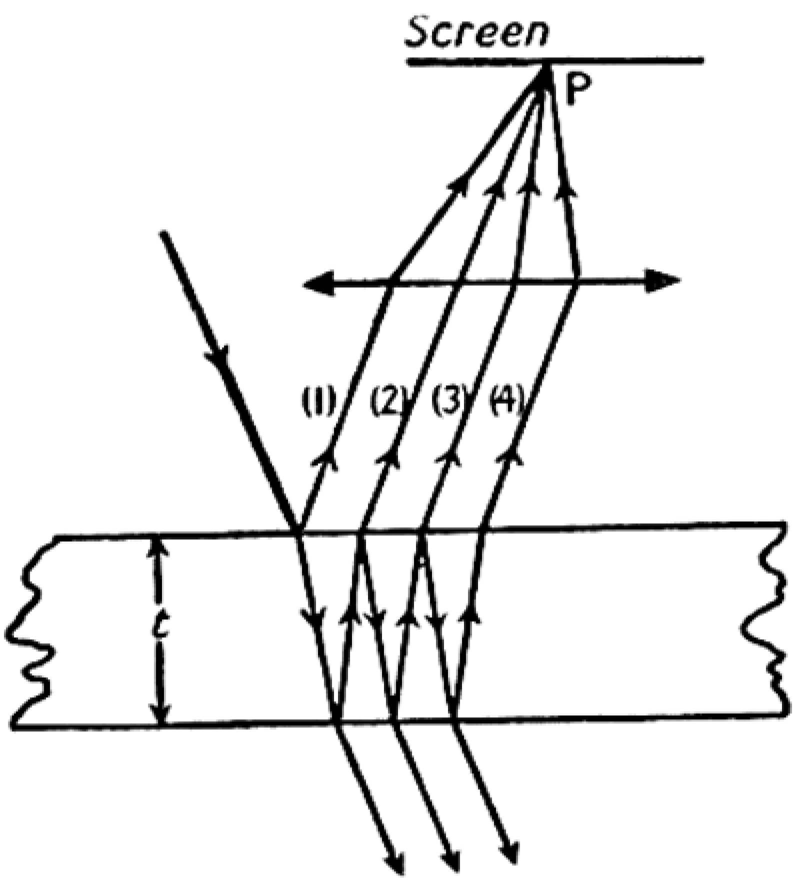


FIG. 718

It will be noticed that in Fig. 718 additional rays after the first pair (which we have already discussed) are shown emerging from the plate after successive reflections. These additional rays will also be brought to a focus at P and will take part in the interference there. The type of reflection which introduces the phase change referred to above occurs only once, namely at A, so that the path difference between (2) and (3), (3) and (4), etc. is simply $2nt \cos r$.

Suppose that the path difference between rays (1) and (2), namely $\left\{(2nt \cos r) - \frac{\lambda}{2}\right\}$, is equal to an odd number of half wave-lengths, so that these two rays tend to cancel each other at P. The condition is expressed mathematically by the equation

$$2nt \cos r = m\lambda$$

where m can be 0, 1, 2, 3 . . . The path difference between adjacent rays like (2), (3), (4), etc. is therefore $m\lambda$, which means they arrive together in phase with each other at P and so reinforce each other at that point. The intensity of ray (1) will be considerably greater than that of (2), which is greater than (3) and so on, because each reflection is only partial. The combined amplitudes of the waves which have travelled along the rays (2), (3), (4) . . . can be shown theoretically to be equal to the amplitude of wave (1) and, since this latter wave is out of phase with all the rest, there will be complete cancellation at P.

Suppose now that the inclination of the incident ray is such as to satisfy the equation

$$2nt \cos r = (m + \tfrac{1}{2})\lambda$$

where $m = 0, 1, 2, 3 \ldots$ In this case the effective path difference between (1) and (2), namely $\left\{(2nt \cos r) - \frac{\lambda}{2}\right\}$, is equal to $m\lambda$, so that these two rays reinforce each other, but the path difference between (2) and (3), (3) and (4), etc. is equal to an odd number of half wave-lengths. Therefore (3), (5), (7), etc. reduce the amplitude of the disturbance at P, being out of phase with (1) and (2), while (4), (6), (8), etc. reinforce (1) and (2). Ray (1) is by far the strongest, and since it is, as we have seen, reinforced by (2), (4), (6), etc., which are together stronger than (3), (5), (7), the result is a maximum illumination of the screen.

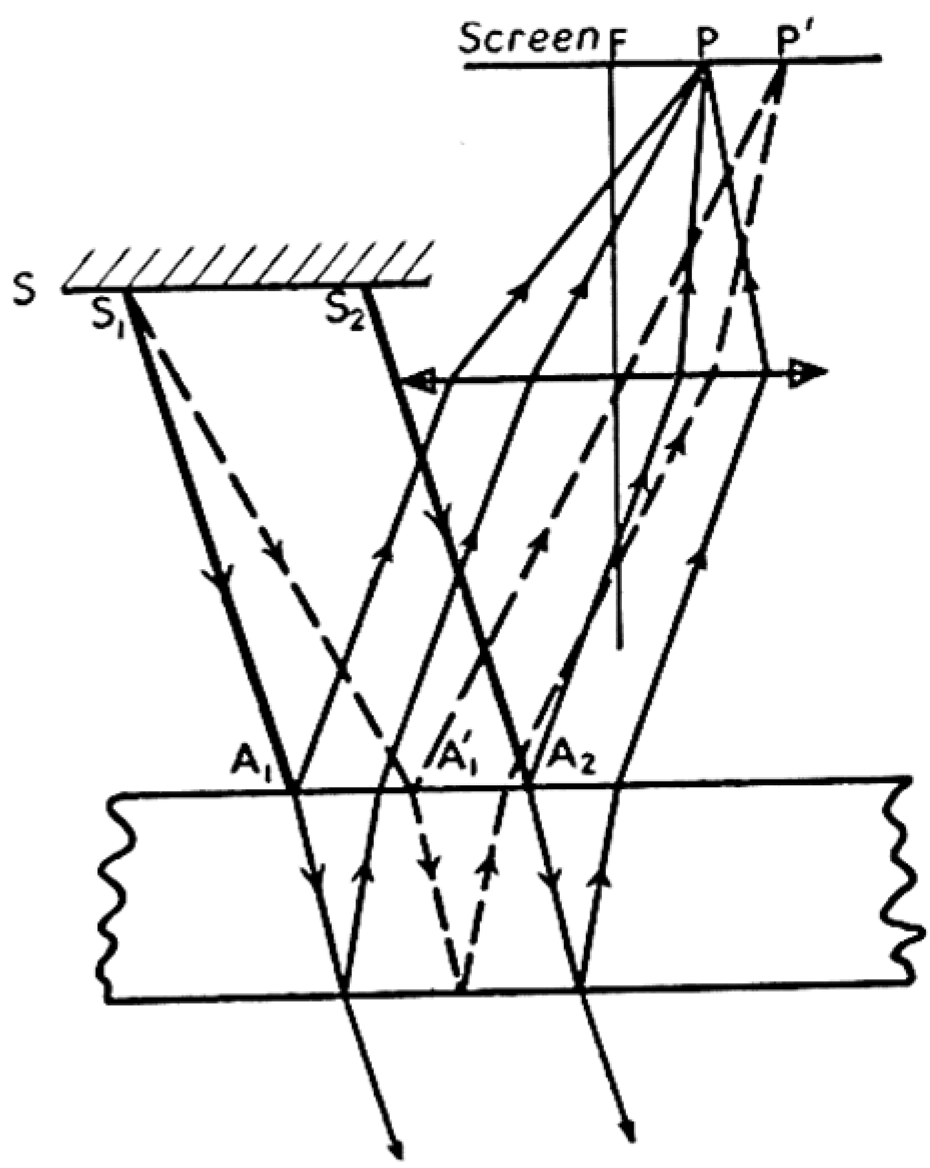


FIG. 719

Let us imagine that a monochromatic source of finite size (a so-called extended source) such as S (Fig. 719) is placed in front of the plate. A ray S_1A_1 from a point S_1 on the source is split up into a series of parallel rays by the plate (only the first two rays of the series are shown), and these are brought to a focus at P in the focal plane of the lens whose axis is perpendicular to the sides of the plate. By the principles of geometrical optics, other rays which are brought to a focus at P must be parallel to the first pair before entering the lens, and therefore must be derived from an incident ray such as S_2A_2 which is parallel to S_1A_1 but originates from another point (S_2) on the source. In fact all rays coming from the extended source which are parallel to the direction S_1A_1 are brought to a focus at P, where the rays from any given point on the source interfere with each other but not with those from other points because they are not coherent sources. The path difference is the same for each of the parallel rays coming from the source, so that if one of them gives a maximum or a minimum at the common focus P, so do they all. A ray like S_1A_1' striking the plate at a different angle gives rise to a series of rays which are brought to a focus by the lens at another point P′ in the focal plane. All other incident rays parallel to S_1A_1' are also brought to P′ where a maximum or a minimum may be produced according to the path difference which, on account of the different value of r, is not the same as for the rays parallel to S_1A_1. Thus all the rays incident on the plate from the extended source are, so to speak, sorted out according to their inclination to the surface of the plate. Moreover, the sorting is made apparent by the fact

that those inclinations which satisfy the equation

$$2nt \cos r = (m + \tfrac{1}{2})\lambda$$

give a bright fringe on the screen, while the inclinations for which

$$2nt \cos r = m\lambda$$

give a dark fringe.

If we imagine that the ray diagram (Fig. 719) is rotated about the principal axis of the lens it is not difficult to realize that, when the source extends at right angles to the plane of the paper, the maxima and minima are actually concentric circular arcs with a common centre at the principal focus F. Each fringe is produced by incident rays which all make the same angle with the plate, and hence the fringes are known as **fringes of equal inclination.**

As already mentioned, the fringes will not be formed on a screen unless a lens is used to bring the interfering rays to a focus. They can, however, be seen by an eye which is directed towards the plate and is focused on infinity so as to bring the parallel rays to a focus on the retina. The eye then replaces the converging lens and screen. Because it is necessary to use some converging system focused on infinity such as a lens and screen, or the eye, in order to bring them into existence, the fringes are said to be **localized** at infinity. In this respect they differ from the fringes due to Lloyd's mirrors, Fresnel's biprism, etc., which can be obtained without focusing on any particular place. Such fringes are **non-localized.** As has already been mentioned, interference of exactly the same nature can take place in a "plate" of air between two parallel glass plates. In this case n in the formula for the path difference stands for the refractive index of air. The marring of the fringes by similar fringes formed by multiple reflections in the glass plates themselves can be avoided by making each of the latter slightly wedge-shaped.

The rays shown issuing from the far side of the plate in Figs. 718 and 719 can also be passed through a lens and made to form fringes in its focal plane in a way similar to that described above for the other set of rays. The path difference between two adjacent rays in the emergent beam is again $2nt \cos r$. Fringes formed in this way are used in the **Fabry-Perot interferometer,** the principle of which is illustrated in Fig. 720. The film of air within which the multiple reflections occur is bounded by the inner surfaces of two glass plates G_1 and G_2. The transmitted rays into which any one incident ray is split up are passed through a converging lens L and are brought to a focus at a point

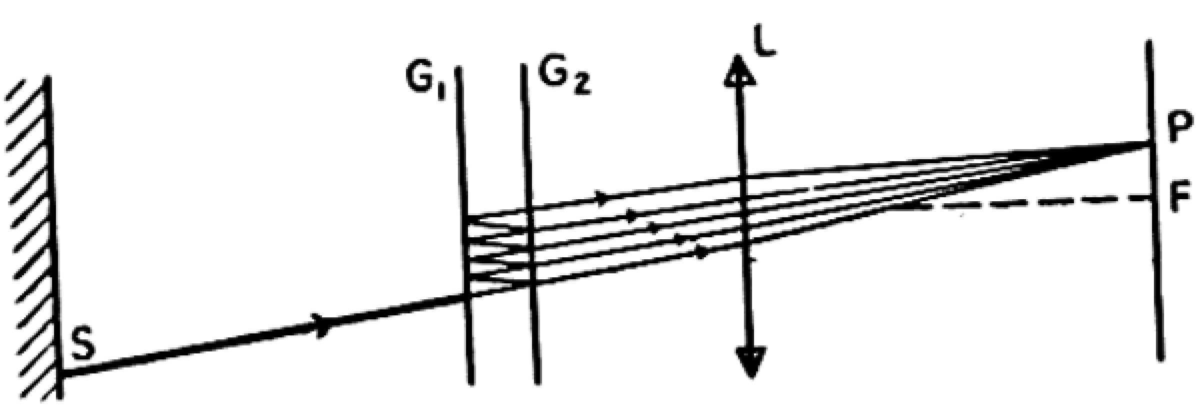


FIG. 720

such as P in its focal plane. All incident rays from the extended source which strike G_1 at a given inclination are brought to the same point in the focal plane, and there is a maximum or a minimum at this point according to the path difference associated with the particular inclination. The fringes are concentric circles with a common centre at the principal focus F, as can be realized by imagining the diagram rotated about the principal axis of the lens. The bright rings in the fringe system can be made very sharp by covering the inner surfaces of G_1 and G_2 with a very thin film of silver, thus increasing the reflecting power and consequently the number of parallel emergent rays taking part in the interference.

In the actual interferometer the separation of the plates can be varied. If sodium light is used and the thickness of the air film is continuously altered by moving one of the plates, the fringe system is seen to become alternately clear and faint. This is due to the fact that, although it can be regarded as monochromatic for some purposes, sodium light actually contains radiations of two different wave-lengths, roughly 5890 Å and 5896 Å. Each wave-length produces its own fringe system, and for certain separations of the plates the bright fringes of one of these coincide with those of the other, thus giving maximum clarity. For other separations the bright fringes of one coincide with the dark fringes of the other and their visibility is at its minimum. The difference of the two wave-lengths can be determined accurately by noting the distance through which one of the plates has to be moved in order to pass from one condition of minimum visibility to the next.

Characteristics of Fringes of Equal Inclination.—Before leaving the subject we shall summarize the characteristics of fringes of equal inclination which have already been mentioned, and shall add one or two more.

1. They are formed by reflection at or transmission through a parallel-sided slab.
2. They are localized at infinity.
3. They are circular in shape.
4. Since the expression $2nt \cos r$ has its *maximum* value when $r=0$, the path difference is greatest at the central fringe, which is due to rays incident normally on the plate.
5. The fact that the differential coefficient of $2nt \cos r$ with respect to r is $-2nt \sin r$ shows that the path difference decreases with increasing inclination, *i.e.* as we go outwards from the centre. Also the rate of decrease becomes greater as r increases, so that the further we go from the centre the closer together the fringes become.
6. Since the rate of change of path difference is proportional to t, the fringes are more crowded together when the distance between the two reflecting surfaces causing the interference is large.

The Use of White Light.—We have seen that when white light is used

in an experiment such as Fresnel's biprism, fringes are visible only in the region where the path difference is zero or very small. For larger path differences the fringe systems due to each colour overlap each other and produce uniform illumination. In general, therefore, a parallel plate will not produce fringes with white light because there is no region of zero path difference.

Consider, however, the action of, say, a very thin oil film on a wet road. An eye looking at such a film receives rays from the sky which have been split up by reflection at the top and bottom of the film in the way we have been discussing. Because t is small, the fringes produced by light of any one wave-length would, as already explained, be broad and widely separated, the path difference varying only slowly with the inclination of the incident rays. At any one inclination, however (as determined by the position of the eye relative to the region of the film looked at), there will be only certain wave-lengths which give a maximum, *i.e.* a bright fringe. Other wave-lengths will suffer some diminution of intensity, and some wave-lengths, which satisfy the condition for a dark fringe at this particular inclination, will be absent altogether. Therefore the light received by the eye differs from white light in that all the component colours except certain ones are diminished in intensity by varying amounts. This is the cause of the well-known colours which are visible in such thin films. The colour varies with the thickness of the film and with the position of the eye, which fixes the value of r.

Interference due to a Wedge.—We now discuss the effect of using a plate of material with plane sides slightly inclined to each other at a small angle—in other words a thin wedge of small angle. We consider first how the geometrical optics of the system differ from those for a parallel plate. Consider, as in the case of the parallel plate, a single ray SA (Fig. 721) incident at any inclination on one face of a wedge of material of refractive index n, the plane of incidence being perpendicular to the line of intersection of the two faces of the wedge. The ray AD is produced at the first reflection while the remainder of the incident light enters the plate and, after passing to and fro once, gives rise to the ray CE. There are other rays as well due to multiple reflections, but only the first two are shown. Since they come from the same point on the source, AD and CE will produce interference if they are brought together by a converging lens to a point P as shown in the drawing. The amount of light at P will depend upon the path difference between the two rays.

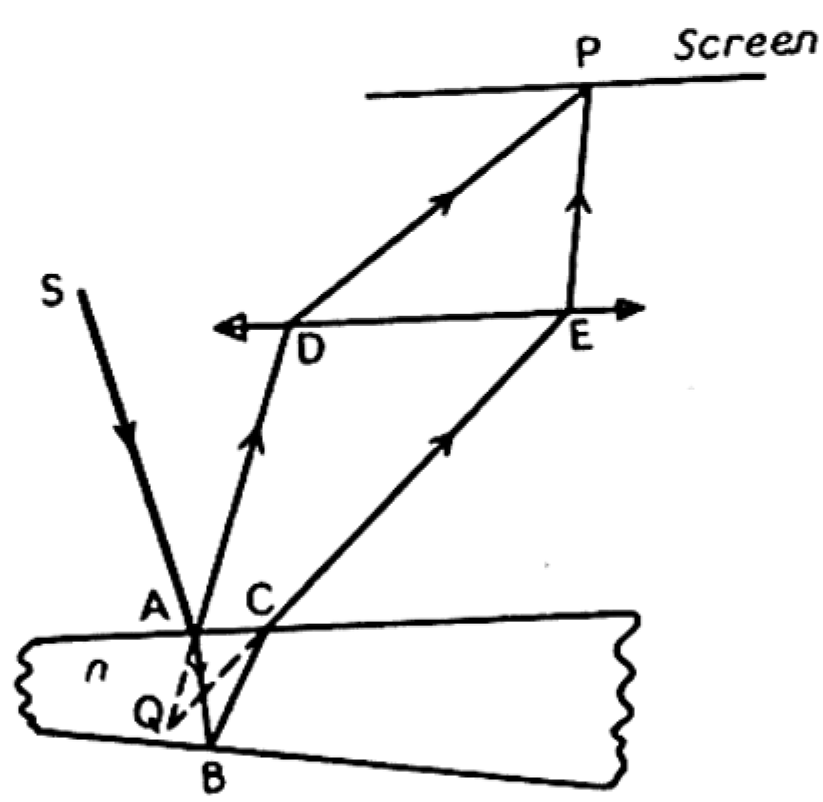


FIG. 721

The distinction between this case and the parallel plate which we have previously discussed is that in the case of the wedge the interfering rays are not parallel. If they are received by an eye, this must be focused on the point Q, and not on infinity, in order that the interfering rays shall arrive at one and the same point on the retina. If the screen is placed at such a distance from the lens as to receive images of points on the top face of the wedge (Fig. 722), then the rays taking part in the interference on the screen will not be those shown in Fig. 721 but the rays SABCE and SCG, both originating from the same point S on the source. The point P is the image of C formed by the lens. Instead of using a lens, the eye may be focused on the top face of the wedge, but more usually, and in order to measure the fringes, a microscope focused on the same surface is used. The fringes seen are said to be localized on the top surface of the wedge since the device used for observing them is focused on that plane. There are many possible ways in which interference can occur, *e.g.* as in Fig. 723, and for each case the position in which the fringes are localized is different. Which of the various possible fringe systems is observed depends upon the focusing of the eye or microscope. However, the fringes are usually seen most clearly when the eye or microscope is focused on the surface (Fig. 722), and we proceed to discuss this system.

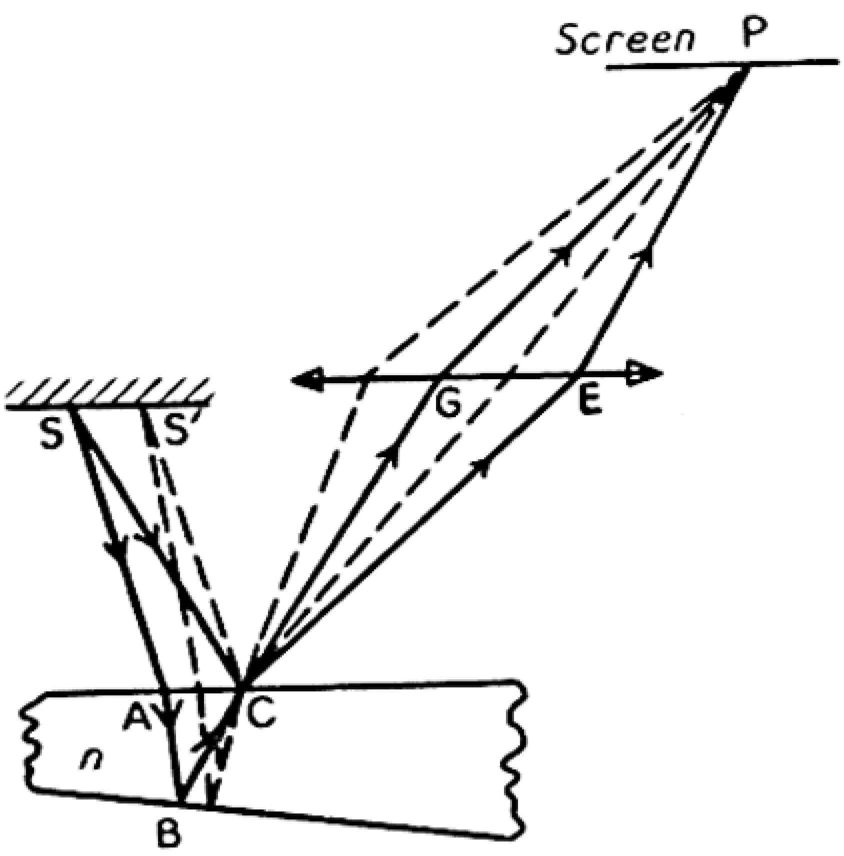


FIG. 722

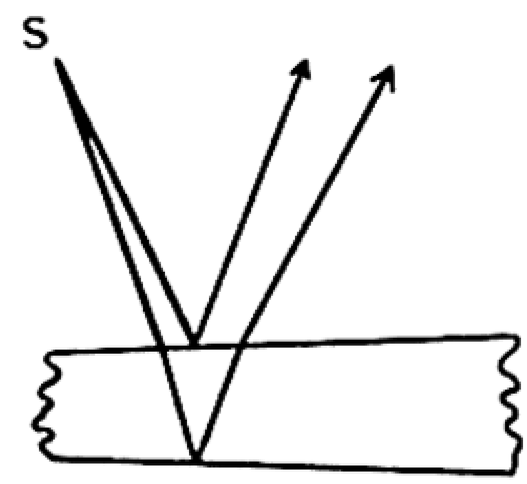


FIG. 723

It can be shown that when the angle of the wedge is very small the path difference at any place where the thickness is t can be taken to be the same as it would be if the plate had parallel sides. It is therefore equal to $\left\{(2nt\cos r)-\frac{\lambda}{2}\right\}$, the term $\frac{\lambda}{2}$ being present on account of the phase change on reflection at the top surface. Referring to Fig. 722, consider a pair of rays from a point S′ on the source other than S which also intersect at C. These rays, which are dotted in Fig. 722, will be brought to a focus at the same point P, but the path difference between them will not be the same as for those coming from S on account of the different value of r. Therefore it appears that it would not be possible with an extended source to

obtain clear interference fringes, because the maxima produced by one point on the source would be superposed on the minima due to another. There are two factors, however, which can be made to reduce the effect of this superposition. The first is the thickness of the wedge, t. When this is small, the rate of change of path difference with r (and therefore with the position of the point on the source) is small, as already explained (page 933), so that the thinner the wedge the more closely will the fringes due to S and S′ coincide, and the clearer the resulting fringes will be. Secondly, if the aperture of the single lens or microscope used for receiving the interfering rays is small, then the rays entering it are confined to a small range of inclinations and their separate fringe systems coincide more closely. The effective aperture of the eye held at a reasonable distance is quite small enough to render clear fringes visible, and their clarity can be increased by holding a pinhole in front of the eye.

Thus, when the foregoing conditions are fulfilled, the wedge is seen to be crossed by a series of alternate light and dark fringes, and the change of path difference between one fringe and the next is due to the change of t as the point C moves along the wedge. Each fringe, therefore, represents the locus of a given value of t, and for this reason the fringes are called **fringes of equal thickness.** An experiment making use of them is described in the next paragraph.

Determination of the Angle of a Wedge.—An air wedge is used for this experiment. It can be made by placing two thin slips of glass face to face, separating them at one end by a thin piece of metal foil and then clamping them together. The air wedge W (Fig. 724) is then laid horizontally on a black surface, and a travelling microscope is arranged so as to look vertically downward on the wedge with its linear scale parallel to the length of the wedge. A small clear glass plate G is placed between the microscope and the wedge with its plane at 45° to the vertical, so that light from a monochromatic source S, such as a sodium flame, is directed downwards on to W. If the microscope is focused on the wedge, the rays which are brought to a focus at the centre of the crosswires all come from one particular point on the wedge, where the thickness is, say, t, and they are all contained within the narrow cone which enters the microscope from that point. The value of cos r for each ray is therefore sufficiently near unity to make the path difference equal to $2nt$ for all the rays coming to a focus at the centre of the crosswires, and indeed for all rays brought to a focus on a line through the centre parallel to the edge of the

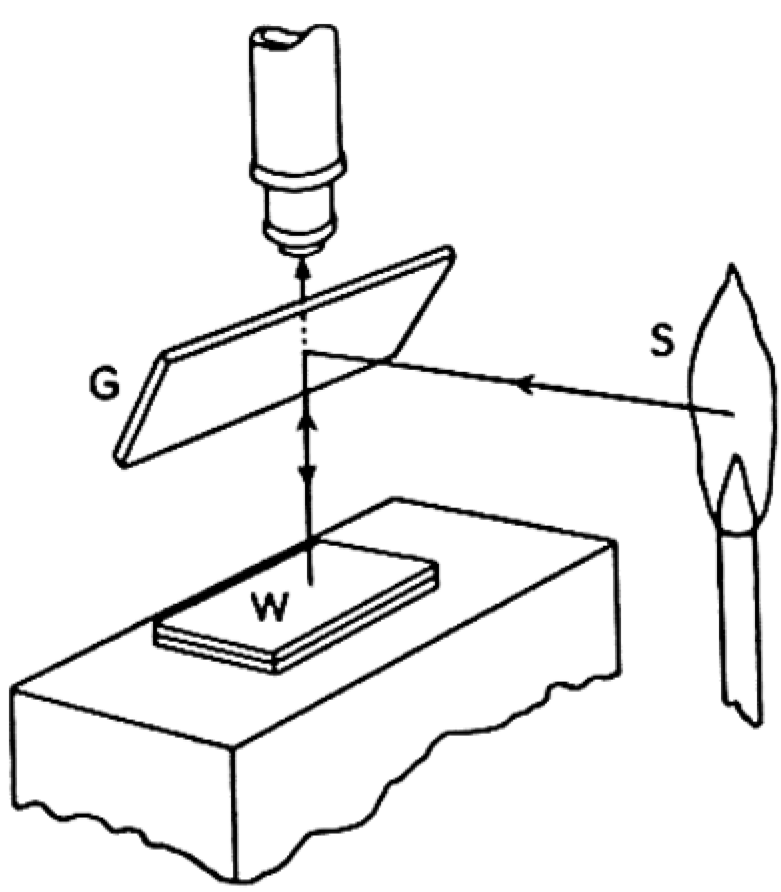


FIG. 724

wedge. Remembering the change of phase on reflection at the bottom of the air wedge and the fact that n is approximately unity, we can say that the fringe on the crosswires is light or dark according to whether $2t$ is an odd or even number of half wave-lengths. Therefore, if the reading of the microscope scale is taken when, say, a light fringe is on the crosswire and the microscope is then moved until the mth light fringe from the first is on the crosswire and the reading is again taken, we know that the difference of the two microscope readings is the distance (d) along the wedge in which the path difference $2t$ changes by $m\lambda$ where λ is the wave-length of the light. The thickness of the wedge t therefore changes by $\frac{m\lambda}{2}$ in a distance d, and consequently the angle of the wedge must be equal to $\frac{m\lambda}{2d}$ radians. The determination of the angle of an air wedge by this method is a common laboratory experiment. The fact that the wedge is bounded by two parallel-sided glass plates does not affect the above argument.

Newton's Rings.—In order to produce the type of interference fringes known as Newton's rings, a plano-convex spherical lens (or possibly a double convex lens) is placed on a flat glass plate. This forms an air film between the lens and the plate, the thickness of which increases outwards from zero at the point of contact of the two surfaces. Evidently the thickness of the film is the same at all points on any circle having the point of contact as centre. Therefore if the eye or a microscope is focused on the film and receives monochromatic light reflected from or transmitted by the surfaces bounding the film, then a series of circular fringes is seen. They are fringes of equal thickness of exactly the same nature as those formed by a wedge with plane surfaces.

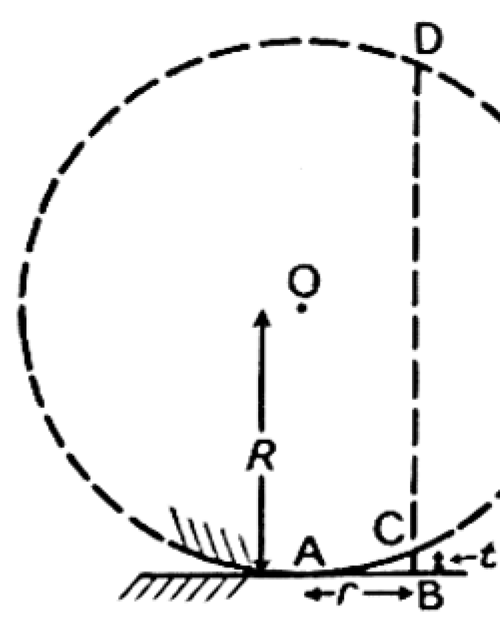


FIG. 725

The radius of any circular fringe can be related to the radius of the spherical surface in the following way. In Fig. 725 the sphere of radius R of which the spherical glass surface is a part is shown completed. A is the point of contact of the curved and plane surfaces, and the thickness BC of the air film is t at a distance r from A. The line BC is produced to meet the circle again at D, and by a well-known theorem in geometry we have

$$(BC) \times (BD) = (AB)^2$$

which we can write

$$t(2R - t) = r^2$$

Since observable interference only occurs when t is very small, we can

neglect this quantity in comparison with $2R$ and write

$$2tR = r^2$$

or

$$2t = \frac{r^2}{R} \qquad \qquad (1)$$

We can produce and measure the circular fringes in exactly the same way as was described for the case of the wedge (Fig. 724). With this arrangement the incident and reflected light is normal to the flat glass surface, so that at a point where the thickness of the air film is t the path difference due to this thickness is $2t$. A change of phase of π occurs, however, during the reflection at the lower boundary of the air film (as already mentioned on page 927), while there is no phase change at the upper (curved) boundary because here the reflection takes place in the glass. Therefore we should write $\left(2t - \frac{\lambda}{2}\right)$ for the actual path difference between the rays which are reflected at the top and bottom of the air film and subsequently interfere in the plane of the microscope crosswires. Thus there is a bright fringe where

$$2t - \frac{\lambda}{2} = m\lambda$$

i.e. where

$$2t = (m + \tfrac{1}{2})\lambda \qquad \qquad (2)$$

where m is 0, 1, 2, 3 . . .

Dark fringes occur where $\left(2t - \frac{\lambda}{2}\right)$ is an odd number of half wave-lengths, *i.e.* where $2t$ is an even number of half wave-lengths or

$$2t = m\lambda \qquad \qquad (3)$$

where m is 0, 1, 2, 3 . . . From this last relation it will be seen that a *dark* area occurs at the point of contact of the curved and plane surfaces, *i.e.* at the centre of the fringe system, because $t = 0$ at this point. The effective path difference is not zero at this point but $\frac{\lambda}{2}$ on account of the phase difference due to the two types of reflection which the interfering rays undergo.

If we eliminate t from equations (1) and (3) we obtain

$$r_m = \sqrt{mR\lambda}$$

where r_m is the radius of the mth dark ring, the value of m for the central dark spot being zero. For the complete dark ring nearest to this spot, $m = 1$. For the next larger dark ring $m = 2$ and so on.

For the bright rings we have, combining equations (1) and (2),

$$r_m = \sqrt{(m + \tfrac{1}{2})R\lambda} \qquad \qquad (4)$$

where r_m is the radius of the mth bright ring, m being equal to zero for the first bright ring immediately surrounding the central dark spot.

It will be seen that the radii of the dark rings are proportional to the square roots of the numbers 0, 1, 2, 3, etc. Owing to the fact that the thickness of the air film increases at a growing rate from the centre outwards, the rings become more crowded together the larger their radius.

The usual experiment performed with the Newton's rings apparatus set up as indicated in Fig. 724 consists in measuring the diameter of a number of bright rings with a view to calculating R or λ from equation (4). The diameters and not the radii of the rings are measured because of the uncertainty of the positions of the centre. If we insert $d_m(=2r_m)$ for the diameter of a ring and square equation (4) we obtain

$$d_m{}^2 = 4(m + \tfrac{1}{2})R\lambda \quad . \quad . \quad . \quad . \quad (5)$$

Therefore if we measure the diameter of the mth bright ring (m being zero for the innermost bright ring, 1 for the next and so on) and then that of the (m')th bright ring, we have for the latter

$$d_{m'}{}^2 = 4(m' + \tfrac{1}{2})R\lambda$$

Thus, by subtraction,

$$d_{m'}{}^2 - d_m{}^2 = 4(m' - m)R\lambda$$

from which R can be obtained if λ is known ($5{\cdot}89 \times 10^{-5}$ cm. for sodium light). Alternatively λ can be calculated if R is measured independently with a spherometer. In practice, a large number of pairs of rings are measured for the same value of $(m' - m)$ and the mean value of the difference of the squares of the diameters found. An alternative method of treating the experiment is to plot a graph of $d_m{}^2$ against m. It is a straight line with a slope of $4R\lambda$, as can be seen from equation (5).

The Coating of Lens Surfaces.—This is an application of interference which reduces the loss of light caused by reflection at the surface of a lens. It consists in depositing on the lens surface a film of transparent substance, of refractive index less than that of the material of the lens, to a thickness equal to one-quarter of the wave-length of light in the film. When light falls (normally) on the film it is reflected partly from the front surface of the film and partly from the surface in contact with the lens. In each case the light is incident on the reflecting surface from the side on which the medium has the lower refractive index, so that both rays suffer the same phase change, and the path difference between them when they have returned to the air is simply twice the thickness of the film, *i.e.* half a wave-length. Therefore if the refractive index of the film is so chosen as to make both reflections equally strong, there is destructive interference between the two reflected rays, which means that no energy is lost by reflection. The transmitted ray which is reflected twice across the film suffers a total phase change of π, which, together with the path-difference

of half a wave-length due to traversing the film twice, brings it into phase with the directly transmitted ray. The light entering the lens is therefore correspondingly stronger than when no film is present, in which case light is lost by the single reflection.

EXAMPLES LI

1. Give an account of Huygens' wave theory of light, and use it to find the relation between the refractive index and wave velocity.

Explain the condition which must be satisfied in order that effects due to the interference of light waves may be observable. (L.I.)

2. A convex lens, of fairly long focal length, is laid on a plane sheet of glass. Describe and explain what is observed when the system is viewed in monochromatic light reflected approximately normally.

State concisely the measurements you would make in order to determine the wave-length of the light with the help of this arrangement, and show how you would calculate the result.

It is required to reduce the amount of light reflected from the surfaces of a lens by depositing a thin transparent film of fluorite (of refractive index 1·43) upon the optically denser glass. Supposing that no light of wave-length 6000 A.U. is reflected when the incidence is normal, calculate a suitable thickness for the film. (O.H.S.)

3. Describe and account for the interference fringes produced when an air film which is enclosed between two plane parallel glass plates inclined at a small angle is viewed normally by monochromatic reflected light.

How may the angle of the wedge be determined by measuring the spacing of such fringes? Give a diagram of the apparatus used. (J.M.B.H.S.)

Chapter LII

DIFFRACTION OF LIGHT

1. FRESNEL'S TREATMENT

Fresnel's treatment of the propagation of waves is briefly discussed in connection with sound on page 589 (Vol. 3) and on page 872 of this volume, and we now give a somewhat fuller account of it as an introduction to the study of the diffraction of light.

In Fig. 726 the dotted outline encloses part of a plane wave-front which is travelling towards P. Let the wave have a definite wave-length λ. The line OP is perpendicular to the wave-front and indicates the direction of propagation of the light. We suppose that every point on the wave-front emits secondary wavelets which are all in the same phase at any given instant and contribute to the total disturbance at P. The separate secondary disturbances originating simultaneously from different points on the wave-front will not, however, reach P simultaneously because the distances which they have to travel are different. In order to find the resultant disturbance at P at any instant it is necessary to add together the secondary disturbances which arrive there *at that instant*. Suppose that the instant chosen is that at which a secondary wavelet emitted by O arrives at P. The disturbance which arrives at P at this same instant from some other point, such as A, is the disturbance which started as a secondary wavelet from A some time before the wavelet which we are considering was given out at O. The difference of time is determined by the difference of the distances AP and OP. For instance, if AP exceeds OP by half a wave-length, the disturbance from A which arrives at P simultaneously with that from O is the one which was emitted from A half a time-period earlier, since this time is required for it to travel the additional half wave-length. In adding together the disturbances which arrive simultaneously at P, it is therefore necessary to take account of the phase differences between them which are determined by their different distances from P. This can be done in effect by treating the wavelets as though they were emitted simultaneously from the wave-front (thereby regarding them all as having the same phase, *e.g.* all maximum displacements), and then taking into account the phase differences which

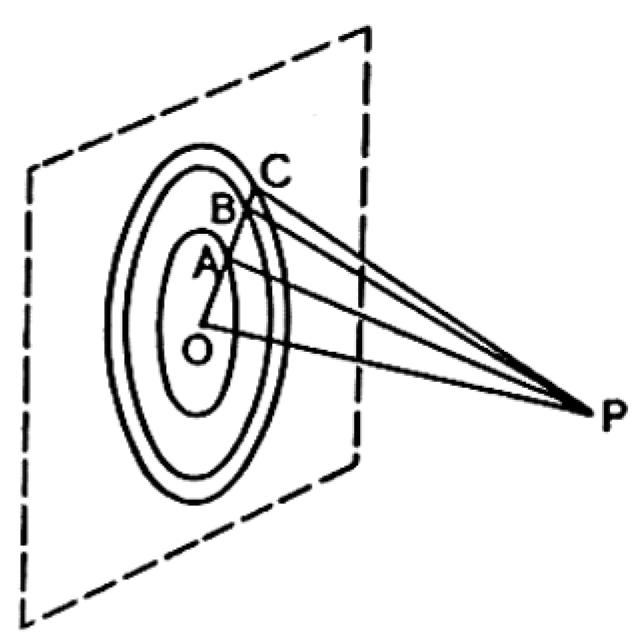


FIG. 726

they would possess on arrival at P as a result of the different distances they have travelled. Thus the wavelets from points on the wave-front whose distances from P differ by a whole number of wave-lengths will reinforce each other at P, while those from points whose distances from P differ by an odd number of half wave-lengths will tend to cancel each other. The calculation of the resultant effect at P, therefore, appears at first sight to be very complicated on account of the infinite number of different points on the wave-front. The matter can be greatly simplified, however, by the treatment due to Fresnel, in which the plane wave-front is imagined to be divided into zones.

With centre O let a series of circles (Fig. 726) be drawn in the plane of the wave-front, their radii being such that, if A is a point on the smallest circle, B a point on the next, C a point on the next and so on, then

$$AP = OP + \frac{\lambda}{2}$$

$$BP = AP + \frac{\lambda}{2}$$

$$CP = BP + \frac{\lambda}{2}$$

The smallest circle and the annular areas between each pair of adjacent circumferences are called **half-wave** or **half-period zones.** For any chosen point in any zone it is possible to find a point in the adjacent zone such that the difference of path from the two points to P is $\lambda/2$. This means that the wavelets from the two points tend to cancel each other at P. But there will not be complete cancellation between the total effects of two adjacent zones unless the total disturbances from each are equal as well as opposite. The intensity of the disturbance at P due to any one zone depends upon the area of the zone, its distance from P, and the angle which the line joining the zone to P makes with OP. This angle is called the "obliquity." The greater the obliquity, the smaller is the contribution made by the zone to the disturbance at P, since the other two effects—area and distance from P—can be shown to cancel each other.

In estimating the total effect at P of the wavelets from all the zones, it can be argued, with some justification, that we can suppose that the effect of any one zone is completely cancelled by half the effect due to the inner adjacent zone and half the effect from the outer adjacent zone. If this principle of cancellation is adopted, it evidently leads us to the conclusion that the resultant disturbance at P due to the whole wave-front is equal to half the disturbance due to the innermost zone *plus* half that due to the outermost zone. It follows that, for a wave-front which is wide compared with the distance of P from it, the disturbance at P is equal to

half the effect of the central zone, because the large obliquity of the outermost zone renders its effect negligible.

The significance of this deduction can be appreciated when the size of the first zone is calculated for any particular case. In Fig. 726, remembering that $AP = OP + \frac{\lambda}{2}$, we have

$$\begin{aligned} AO^2 &= AP^2 - OP^2 \\ &= \left(OP + \frac{\lambda}{2}\right)^2 - OP^2 \\ &= \lambda . OP + \frac{\lambda^2}{4} \\ &= \lambda . OP \quad \text{approximately} \end{aligned}$$

if λ is small compared with OP so that the term $\frac{\lambda^2}{4}$ can be neglected. It will be noticed that for a given wave-length the radius of the zone is approximately proportional to the square root of OP.

Let $OP = 100$ cm., and consider the case of light waves, for which we can assume λ to be equal to 6×10^{-5} cm. For this case we can use the approximate formula, and we have

$$\begin{aligned} AO &= \sqrt{\lambda . OP} \\ &= \sqrt{6 \times 10^{-3}} \\ &= 8 \times 10^{-2} \text{ cm.} \quad \text{roughly} \end{aligned}$$

If the waves are coming from a point source S (Fig. 727) and are spherical, the zones for a point P can be drawn in the same way as for a plane wave as indicated in the figure. For a fairly distant source, the radius of the first zone for a point 100 cm. ahead of the wave-front will be of the same order of magnitude as that for a plane wave-front, which has been calculated above.

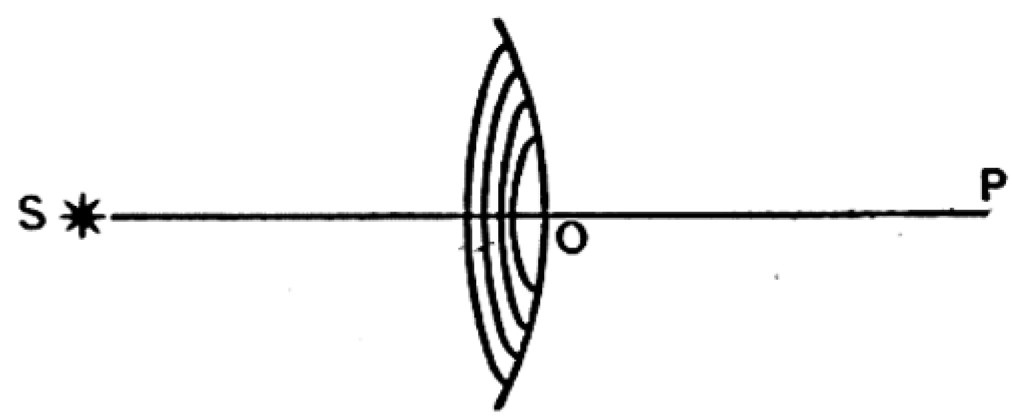


FIG. 727

Since the total effect of the wave-front at P is approximately equal to half the contribution of the central zone, we can regard the effectiveness of the wave-front in respect of the propagation of energy to the point P as being confined to the small area, of the order of magnitude calculated above, situated around the straight line joining P to S. At first sight, the idea of spherical secondary wavelets originating from every point on the wave-front would seem to suggest that the whole wave-front would

transmit energy to any one point, such as P, in front of it. This is indeed the case, of course, but the destructive interference between the secondary wavelets arriving at P can, as we have seen, be represented as rendering ineffective the whole wave-front outside the first zone. In this sense we can say that the wave theory of light gives an indication of approximately rectilinear propagation.

The rectilinear propagation of light is, of course, common knowledge. Sound waves, on the other hand, are able to travel round corners, that is to say, sound "shadows" are normally much less sharp and complete than those due to light. The reason for this contrast lies in the very much longer wave-length of sound waves. By way of illustration we may take this wave-length to be 100 cm., in which case $\lambda/2$ is 50 cm., and the radius of the central zone of a plane wave-front with respect to a point situated 100 cm. in front of it is $\sqrt{150^2-100^2}$ or about 110 cm. (see page 943). This, therefore, is the radius of the zone which is effective in sending the sound to a point 100 cm. ahead of the wave-front as compared with a corresponding radius of less than 1 mm. in the case of light. With such a large effective area of a sound wave-front the propagation cannot be regarded as rectilinear.

If light energy travelled in straight lines, a point source of light would give rise to perfectly sharp shadows, the shape and size of which would be determined by imagining straight lines to be drawn from the source to the edges of the obstacle and produced to the screen. Such a shadow is called a "geometrical" shadow. We see that the matter is much more complicated, however, when we realize that the obstacle obscures only part of the wave-front and that any point beyond the obstacle, whether inside or outside the geometrical shadow, receives light from the unobscured portion of the front. Thus it appears possible that there will be some illumination on the screen within the region of the geometrical shadow and that the edges of the shadow will not be sharp. This is indeed what happens, and various simple cases of the diffraction of light round obstacles are considered in the next section of this chapter. It must be remembered however that, on account of the comparatively short wave-length of light, the diffraction effects we are about to describe are small-scale effects and are not normally observed. For the large-scale effects of light the principle of rectilinear propagation and the conception of light energy travelling along rays is quite adequate.

2. PARTICULAR CASES OF DIFFRACTION *

Circular Obstacle.—In order to demonstrate the diffraction of light into the shadow of a circular obstacle, a small circular hole S (Fig. 728 (i)) in a screen is strongly illuminated with monochromatic light (*e.g.* a sodium-

* With the possible exception of the diffraction grating, the topics dealt with in this section may be beyond the requirements of some readers.

vapour lamp) so as to act as a point source. A small circular disc or ball-bearing is placed a metre or so in front of the hole and casts a shadow on a screen placed about the same distance beyond the obstacle. The most striking result of the experiment is the existence of a bright spot on the screen at P, the centre of the shadow. The phenomenon is explained in terms of Fresnel's zones as follows.

We consider the wave-front of the monochromatic light from the pinhole S at the instant when it is passing over the obstacle. Since the latter is a considerable distance away from the source, the wave-front can be regarded as being practically plane. We now imagine that circles concentric with the obstacle are described on the wave-front as shown in Fig. 728 (ii), in which the shaded inner circle represents the obstacle as seen from P. The radius of the first circle beyond the edge of the obstacle is such that the distance from points on its circumference to P exceeds the distance from points on the edge of the obstacle to P by $\lambda/2$, where λ is the wave-length of the light. Similarly points on the second circle are $\lambda/2$ further from P than points on the first circle, and so on. Thus the circles divide the wave-front into half-period zones. There is no essential difference between the case of the unobstructed wave-front and the one now under discussion in which the central portion of the wave-front is obstructed by the small circular obstacle. Owing to destructive interference the effect at P of the unobscured part of the wave-front is equivalent to half the effect of the zone immediately surrounding the obstacle. Therefore a perceptible amount of light reaches P, provided the size of the object is small compared with its distance from the source and screen. At points other than P in the shadow there may be destructive interference or reinforcement, *i.e.* either darkness or light, according to the phase relationships of the light reaching the point in question from the various parts of the wave-front. As might be expected from considerations of symmetry, light and dark *rings* are seen surrounding the central bright spot. When Fresnel first put forward his theory of half-period zones it was objected to on the very grounds that it implied the existence of light in the centre of the shadow of a circular obstacle—an effect afterwards demonstrated by Arago.

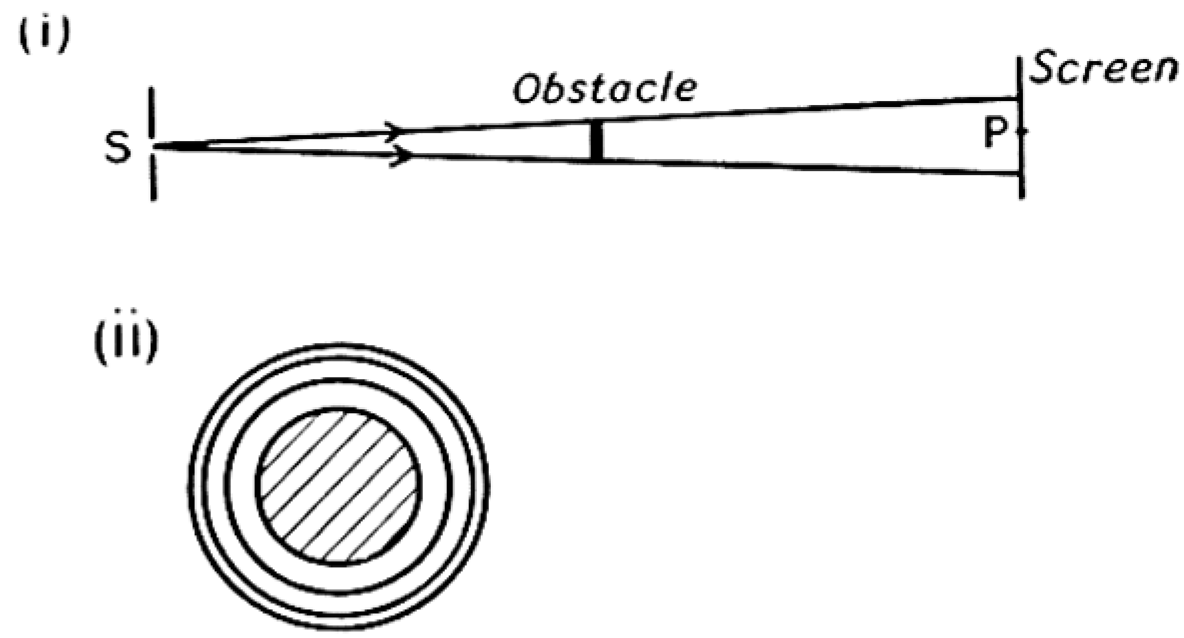


FIG. 728

Circular Aperture.—The arrangement which we discuss under this heading is shown in Fig. 729 (i). Monochromatic light from a small source S passes through a circular aperture (centre O) in a screen and then

falls on another screen. According to the principle of rectilinear propagation a cone of rays would be formed which would cause a uniform circular patch of light with sharp edges on the screen. In order to discuss possible diffraction effects at the point P on SO produced, we construct half-period zones on the wave-front coming through the aperture as indicated in Fig. 729 (ii). These zones are exactly the same as if the wave-front were unobstructed except that their number is limited by the edge of the aperture. Suppose that we keep all other distances constant and we consider the effect of the size of the aperture on the illumination at P.

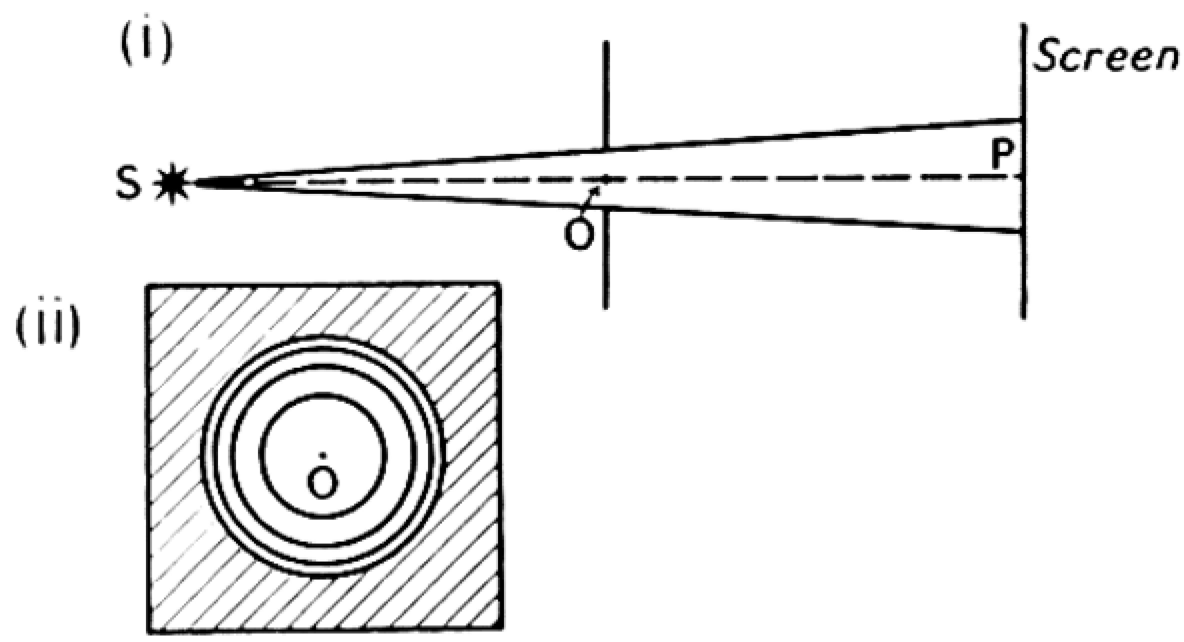


FIG. 729

When the aperture is comparatively large the wave-front is practically unobscured and, as explained on page 942, the illumination at P is equivalent to half the effect of the innermost zone, the zone nearest the edge of the aperture being practically ineffective. Thus there will be the same amount of light at P as if the screen containing the aperture were not interposed.

Next suppose that the aperture is so small as to cause all but the innermost zone to be covered. Obviously the effect at P is now the *entire* effect of this zone, since there are no other zones to oppose its effect as there are when the whole wave-front is unobscured. We therefore reach the surprising conclusion that, by interposing a screen with a small aperture of the correct size, we actually increase the illumination at P due to S above what it would be if the wave-front were entirely unobstructed. In fact there is a fourfold increase in the intensity of light at P because the *amplitude* of the disturbance is doubled (being in the unobstructed case due to half the central zone and in the other case due to the whole zone) and the energy of the radiation at P is proportional to the square of the amplitude (page 562). If the distance SO is large enough to allow us to regard the wave-front passing through the aperture as plane, and if we take the distance OP as being 100 cm. and the wave-length of the light as 6×10^{-5} cm., then, as explained on page 943, the radius of the innermost zone is about 0·8 mm. Therefore an aperture of this size will have the effect just described.

Next suppose the aperture is enlarged (actually to a radius of about 1·1 mm.) so that it limits the wave-front to the first *two* zones. The effects of these two zones, being out of phase with each other, cancel each other at P and produce a *dark* spot. A further increase in the size of the aperture to include three zones restores the bright spot at P, because now

the second zone can be regarded as cancelling half the effect of the first zone and half the effect of the third, thus leaving half the joint effects of the first and third zones which are in phase (*cf.* the treatment of the whole wave-front on page 942). With four zones P again becomes dark and so on. The effect on the intensity at P diminishes, however, as the number of zones increases, because the effect of the outermost zone decreases.

The same effect of periodically varying illumination at P can be obtained by keeping the size of the aperture constant and moving the screen towards and away from the aperture. This has the effect of changing the sizes of the zones and therefore the number of zones contained in the aperture.

The illumination on the screen fluctuates as we go outwards along the screen from P and, as might be expected, P is in fact surrounded by alternate light and dark rings.

Diffraction due to a Straight Edge.—We shall discuss the diffraction effects seen in the immediate neighbourhood of the edge of a shadow by considering the case of a straight edge in front of and parallel to a slit through which monochromatic light is passing. Fig. 730 is a perspective schematic diagram showing a position of the wave-front, which, it is easily understood, is cylindrical in shape with the slit SS′ as axis. We now consider what effect there would be on a screen placed beyond it if the wave-front is partially obstructed by a second screen having a straight edge parallel to the slit. Without proving it rigorously, we can state that the symmetry of the arrangement suggests that whatever pattern is produced will be in the form of straight lines parallel to the slit. Therefore, instead of considering the effect of the wave-front at a point, as we have done when discussing spherical and plane waves, we can, in this case, consider a *line* such as PP′ on the screen, and we can divide the wave-front into half-period strips with respect to PP′ as shown in Fig. 730. The nearest line to PP′ on the wave-front is OO′, and the first strip on one side of OO′ extends as far as AA′, which is a line whose distance from PP′ exceeds that of OO′ from PP′ by $\lambda/2$. The next strip on the wave-front finishes at BB′ and so on, and there are exactly similar strips on the other side of OO′.

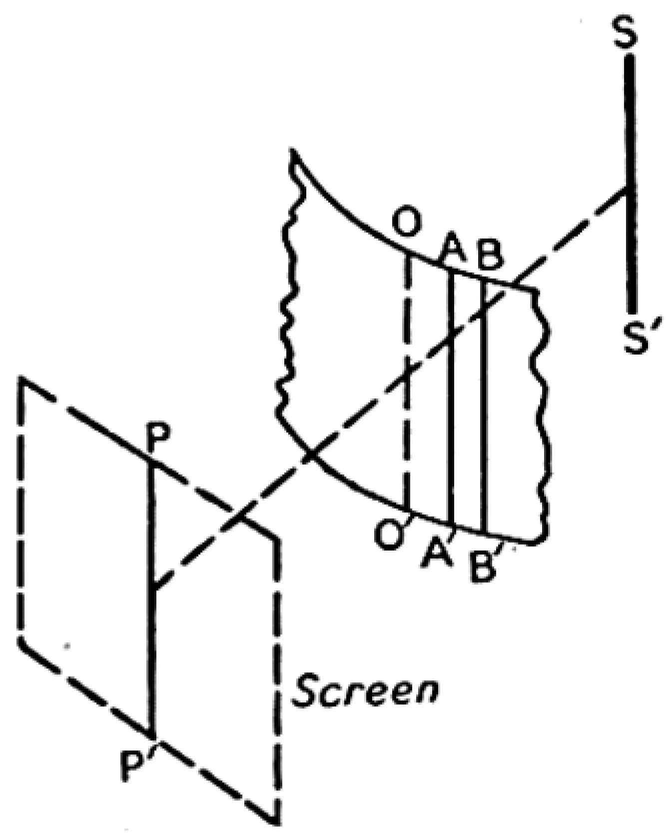


FIG. 730

In the absence of any obstruction of the wave-front there are a large number of half-period strips on each side of OO′, and the effect at PP′ of the strips on one side of OO′ is due mainly to the first strip (OO′A′A in Fig. 730), chiefly because it has the largest emitting area. A similar statement applies to the strips on the other side of OO′, so that each of the two sets of strips contributes equally to the disturbance at PP′.

In Fig. 731 (i), in which we are looking down on the arrangement shown in Fig. 730, the wave-front is partially obscured by a screen having a straight edge at O parallel to the slit, so that the point P is on the edge of the geometrical shadow of the obscuring screen. We can then deduce that, since the strips below O are now ineffective, the amplitude of the disturbance at P is simply that due to the unobstructed half of the wave-front and is therefore half the disturbance due to the whole wave-front.

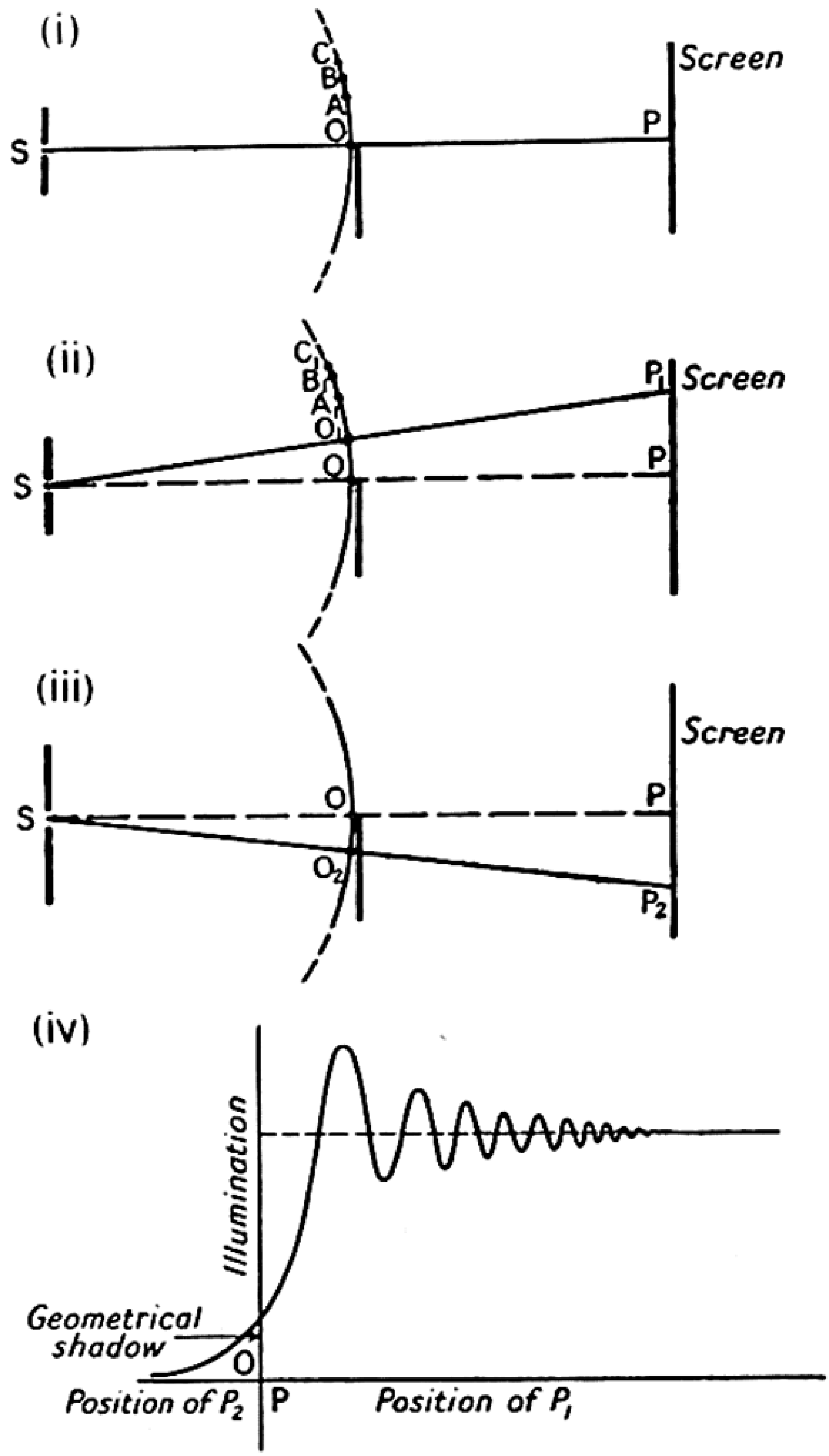


FIG. 731

We now seek to explain the graph drawn in Fig. 731 (iv), which shows the distribution of illumination (proportional to the square of the amplitude) on the screen in the neighbourhood of the geometrical shadow, the edge of which (*i.e.* P) is taken as the zero on the horizontal scale. We consider first a point such as P_1 (Fig. 731 (ii)) outside the geometrical shadow, and it is important to realize that as P_1 moves so also do the half-period strips into which the wave-front can be divided with respect to the point P_1. The corresponding pole to P_1 is O_1, and the boundaries of the half-period strips above O_1 are marked as A_1, B_1, etc. Since none of these strips is obscured, they jointly produce at P_1 a disturbance of the same amplitude as exists at P, since this latter is also due to half the wave-front. In addition, the strips between O_1 and the edge O of the obstacle contribute to the disturbance at P_1. As P_1 moves up from P, O_1 moves up from O, and the first strip below O_1 is gradually uncovered, thereby increasing the disturbance at P_1. The first and highest maximum of the graph is reached at a position of P_1 such that the first strip is nearly (but in fact

not quite) completely exposed. This strip is then contributing more to the disturbance at P_1 than the wave-front above O_1.

When P_1 moves further up the screen the second strip below A begins to be exposed, and since, on account of its opposite phase, its effect at P_1 tends to cancel that of the first strip, the illumination on the screen decreases and reaches a minimum value when the second strip is nearly completely exposed. With P_1 moving higher still a second maximum is reached when the first three strips below O_1 are exposed, because the disturbances due to both the first and third strips are in phase with the resultant disturbance from the upper part of the wave-front above O_1, and only the second strip below O_1 tends to diminish the disturbance because of its opposite phase. Without continuing the argument, it will now be realized that there are bands or fringes of fluctuating intensity on the screen *outside* the geometrical shadow as indicated by the graph in Fig. 731 (iv). The following points should be noticed. Firstly, the minima do not represent complete darkness. Secondly, the contrast between consecutive maxima and minima diminishes as we go away from P, because the last strip to be exposed above O becomes narrower and less effective until P_1 is sufficiently far from P for the wave-front between O_1 and O to be effectively as complete as the part above O. Thirdly, the fringes become crowded together as their distance from P increases. This is due to the fact that the strips get narrower as their distance from O_1 increases, so that a correspondingly small movement of P_1 is required in going from one fringe to the next.

Finally, we consider the appearance on the screen in the region below P, *i.e.* in the geometrical shadow. The point O_2 (Fig. 731 (iii)) is the pole for P_2, and the disturbance at P_2 is due to the unobstructed part of the wave-front above O, which can be divided into half-period strips beginning at O such that the distances of the two edges of any one strip from P_2 differ by $\lambda/2$. As P_2 moves down the screen from P the effect of the uncovered part of the wave-front diminishes rapidly but does not fluctuate. This explains the small portion of the graph in Fig. 731 (iv) which lies to the left of the vertical axis, *i.e.* within the geometrical shadow. It is important to realize that although the change of illumination in the immediate neighbourhood of the edge of the geometrical shadow is not abrupt, yet, owing to the smallness of the wave-length of light, the change is practically complete within a distance of a millimetre or so. This is why the lack of sharpness normally escapes unnoticed. It is further masked by the fact (page 698) that in any case all but the smallest light sources give shadows with non-sharp edges (the penumbra). With sound, of wave-length perhaps a million times that of light, the same change of intensity near the edge of the shadow takes place over a correspondingly larger distance. This accounts for the lack of sharpness of sound shadows, which has already been discussed on page 944.

Diffraction by a Narrow Strip.—When a narrow strip (or wire) is set

up parallel to, and some distance from, a slit source, it produces diffraction patterns in he form of straight fringes parallel to the length of the obstacle. The arrangement is shown in Fig. 732. If the strip were sufficiently wide, each of its two straight edges would act independently in the way described in the foregoing paragraphs. The result would be a set of fringes outside the geometrical shadow on each side, together with a narrow region of diminishing illumination extending inwards from each edge of the geometrical shadow. The strip may be made sufficiently narrow, however, to cause the last-mentioned regions to overlap inside the shadow. In this case the two sets of fringes outside the shadow still exist, while a third central set are formed by interference of the light entering the shadow round each edge of the obstacle. In fact, the edges of the obstacle can be regarded as two coherent line sources producing interference in just the same way as the two slit images in the biprism (page 923), and the formula giving the separation of the fringes can be worked out in this way. Thus, as on page 925, we obtain the result that the fringes are equidistant. For a light fringe the distances of the two edges of the obstacle from the fringe must differ by a whole number of wave-lengths, and for a dark fringe the path difference must be an odd number of half wave-lengths. If the width of the strip is d and the distance of the screen from the strip is D, then the path difference for a fringe situated at a distance x from the centre of the shadow is, as on page 924, $\frac{xd}{D}$. Therefore the values of x for which there is a bright fringe are given by the equation

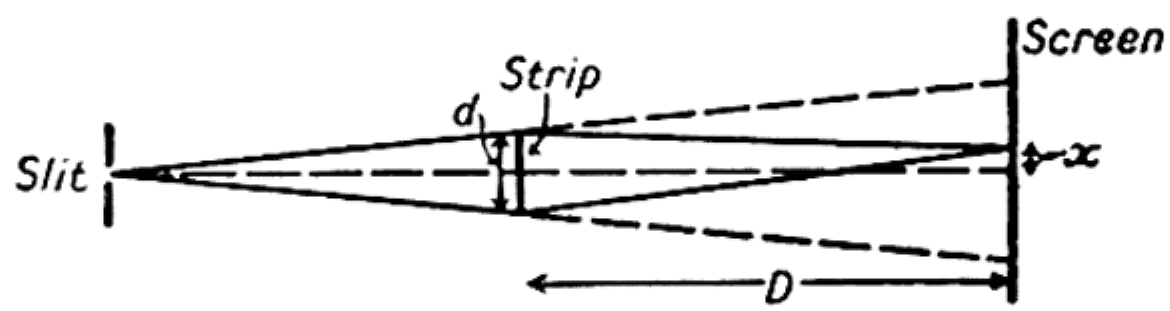


FIG. 732

$$\frac{xd}{D} = m\lambda$$

where m is 0, 1, 2 . . ., while for a dark fringe

$$\frac{xd}{D} = (m + \tfrac{1}{2})\lambda$$

where m is again 0, 1, 2 . . . The distance between consecutive bright (or dark) fringes is equal to the change of x when m changes by unity, *i.e.* to $\frac{\lambda D}{d}$. This can be measured with the help of a micrometer eyepiece as described in connection with the biprism, and a value for λ can then be obtained.

Diffraction due to a Slit.—A diffraction pattern is produced when light from a slit passes through a second slit and then falls on a screen.

The second slit is responsible for the diffraction. The arrangement is evidently the counterpart of the diffraction by a circular aperture of light coming from a small circular hole, and a similar argument can be applied as regards the centre of the patch of light on the screen. The central fringe may be light or dark according to the relation between the width of the slit and the size of zones, and is in all cases flanked by other light and dark fringes corresponding to the rings which are seen when light is diffracted through a circular aperture.

We shall, however, discuss the diffraction of light by a slit from a rather different point of view. In Fig. 733 (i), AB represents the slit, and the

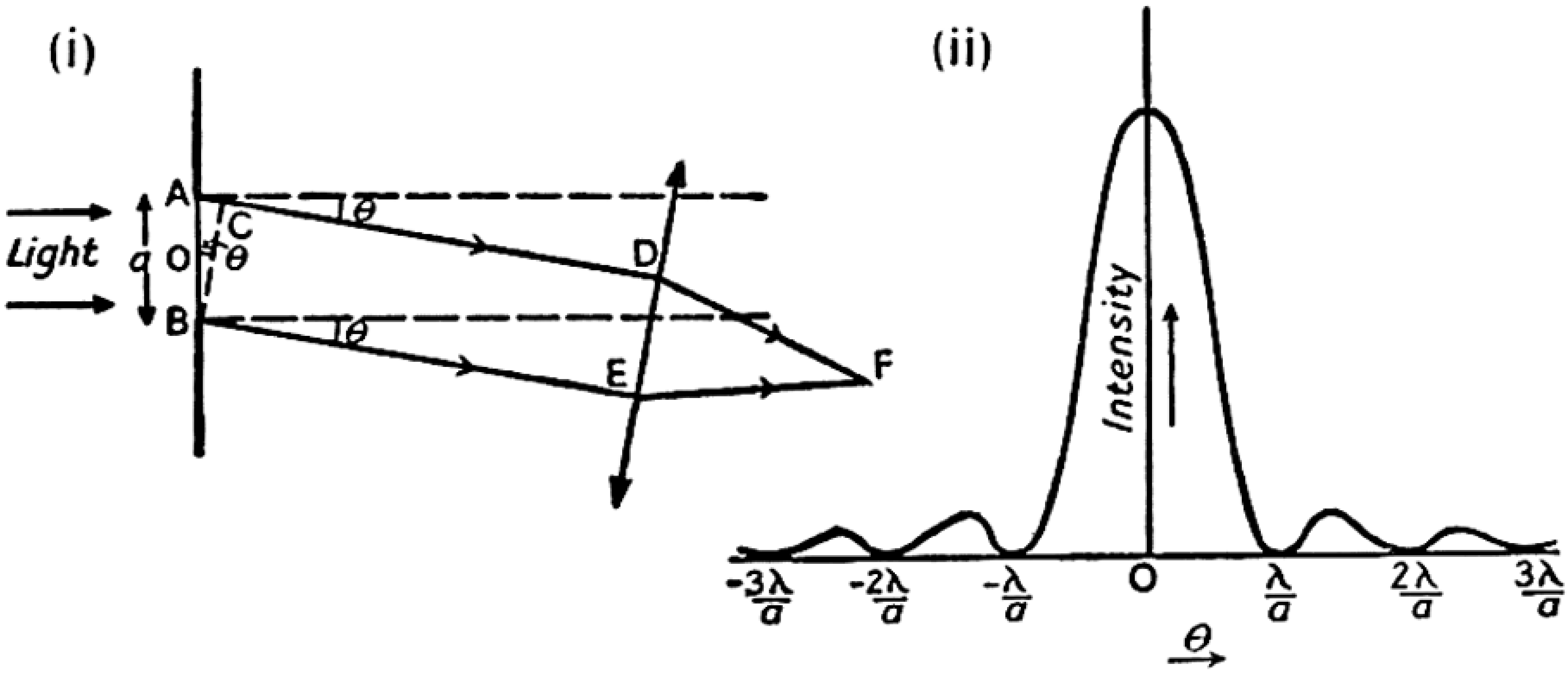


Fig. 733

thin line drawn across it represents a plane wave-front which has reached the slit from a line source which is either very distant or is situated in the focal plane of a converging lens which has therefore rendered the light parallel. According to the Fresnel treatment of diffraction, we suppose that each point on the wave-front AB gives out spherical wavelets. We now consider how the intensity of the light leaving the slit will vary with *direction*. For this purpose, therefore, we consider a direction making an angle θ with that of the incident light. If a converging lens is set up with its axis parallel to this direction, then all the light proceeding from the slit to the lens in this direction will be brought to the principal focus F. A screen may be placed at F or, better still, an eyepiece may be focused on F, thereby forming a telescope focused to receive parallel light.

If BC is drawn perpendicular to the lines AD and BE, then it is clear that the light reaching F from A has travelled a distance AC further than the light which has come from B. Suppose that θ is such as to make AC equal to λ, the wave-length of the light. Let O be the centre of AB. The difference of path between light reaching F from A and light reaching F from O is $\lambda/2$, so that the wavelets from A and O cancel each other when they reach F. The same path difference exists between any point in AO situated a certain distance below A and the corresponding point in OB

situated at the same distance below O. Thus in a direction θ which makes AC equal to λ there will be darkness at F—a straight dark fringe parallel to the slit, in fact. The same effect will be produced if the path difference AC is 2λ, because we can then divide AB into four equal parts and apply the same argument to each pair of quarters. In fact there is a dark fringe at F for those values of θ which make AC equal to a whole number of wave-lengths. If the width of the slit AB is denoted by a, it is evident that

$$AC = a \sin \theta$$

so that the values of θ for a dark fringe are given by the equation

$$\sin \theta = \frac{m\lambda}{a}$$

where m can be 1, 2, 3 . . . (but not zero). Since the pattern is only visible over a small range of θ we can replace $\sin \theta$ by θ and write

$$\theta = \frac{m\lambda}{a} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

As regards the bright fringes, we can say that when θ is zero, *i.e.* when the axis of the lens coincides with the normal to the plane of the slit through O, there will be maximum illumination at F because the wavelets then arrive at F in the same phase. Thus there is a central maximum flanked by the first minimum on each side, for which, as we have seen, the value of θ is given by $\sin \theta = \frac{\lambda}{a}$. There are secondary maxima between (but not midway between) the other minima. We can see roughly how these are formed, because if the path difference AC is, say, $3\lambda/2$ we can imagine the slit to be divided into three equal parts. Then the path difference between points in the top third and corresponding points in the centre third will be $\lambda/2$, so that the effects of these two strips cancel each other at F and the illumination is due to the remaining third.

The graph of intensity against θ is shown in Fig. 733 (ii). The minima are shown to occur at $\theta = \frac{\lambda}{a}, \frac{2\lambda}{a}$, etc., according to the approximate equation (1). The central maximum is twice as wide as the others and is very much more intense. The secondary maxima diminish rapidly in intensity with increasing θ.

It will now be realized that this type of diffraction, which is known as **Fraunhofer diffraction,** differs from the cases we have hitherto discussed. In previous examples we have considered the effect of secondary wavelets at a *point* on a screen, there being no lens. This is called **Fresnel diffraction.** In Fraunhofer diffraction we are concerned with the distribution of light with *direction*, and in order to obtain a set of fringes

it is necessary to use a lens or telescope to focus the parallel light leaving the slit in any given direction.

Returning to the diffraction pattern as represented by the graph in Fig. 733 (ii), suppose that the width of the slit is one-tenth of a millimetre (10^{-2} cm.) and that λ is 6×10^{-5} cm. The first minimum therefore occurs at an angle of $\dfrac{6 \times 10^{-5}}{10^{-2}}$, *i.e.* 6×10^{-3} *radians*, on each side of the direct light passing through the slit. If the light is collected by a lens of 10 cm. focal length, the distance between the central maximum and the first minimum in the focal plane will be $10 \times 6 \times 10^{-3}$, *i.e.* 0·06 cm. or 0·6 mm., and it is easily seen that this separation is inversely proportional to the width of the slit. This calculation shows that the whole diffraction pattern is quite small, and can be seen on the screen at one and the same time if the lens is placed with its axis along the normal to the plane of the slit.

The system depicted in Fig. 733 (i) is equivalent to a telescope viewing a distant line source while a slit is placed in front of the objective. The diffraction fringes indicated in Fig. 733 (ii), and especially the comparatively broad central maximum, are then seen instead of a sharp image of the source. However, if the slit is widened the central band becomes narrower and may eventually *appear* as a sharp image although diffraction effects are never entirely absent. It is evident that when the slit is narrow enough to give a central band broader than the geometrical image of the line source, a similar source placed alongside and sufficiently near the first will produce a second band overlapping the first and the two sources will not be distinguishable in the telescope. It is usual to assume that for equal sources the images are just distinguishable from each other when the central maximum of one pattern falls on the first minimum of the other. The angle subtended by the two sources at the aperture is then λ/a radians (page 952). Sources subtending a smaller angle cannot be resolved, and this critical angle is called the **minimum angle of resolution** of the instrument with the particular slit in place. Telescopes are not, of course, normally used with slits in front of them, and we have only used the slit by way of simplified illustration. The same principle applies, however, when the telescope is used in the ordinary way. The diffraction then occurs at the circular aperture of the objective and, when a point source such as a star is viewed, it appears as a central circular patch of finite size surrounded by dark and light rings. The limit of resolution of two point sources is reached when the centre of the central patch due to one point source falls on the first dark ring due to the other, and it can be shown that the angular separation of the point sources is then $\dfrac{1{\cdot}22\lambda}{d}$, where d is the diameter of the aperture of the objective. This formula therefore gives the minimum angular resolution of the instrument. It can be improved by increasing d as already mentioned on page 883.

A similar principle applies to microscopes, and in this case the resolving power is often increased by reducing λ (page 872).

The eye itself has a limit of resolution which is due to diffraction at the pupil. Theoretically the minimum angular resolution is about 50 seconds of arc, but in fact it is not less than about 1 minute on account of the structure of the retina (page 859).

The Diffraction Grating.—The last case of diffraction which we shall discuss is, in many ways, the most important on account of the precise measurements of wave-length to which it leads. A diffraction grating is, effectively, a series of equidistant parallel slits placed side by side. A grating used for the diffraction of light can be made by ruling on the surface of a plane sheet of glass a very large number of fine parallel equidistant scratches. Many gratings are made with about 14,000 lines to the inch. The thin strips of the original glass surface between the scratches act as slits, and when light falls on one side of the grating each slit emits cylindrical wavelets.

We investigate the performance of a grating in the same way as we did the single slit. That is to say, we inquire in what *directions* the secondary wavelets from the grating will reinforce each other. Fig. 734 is a diagrammatic representation of the arrangement. The lines of the

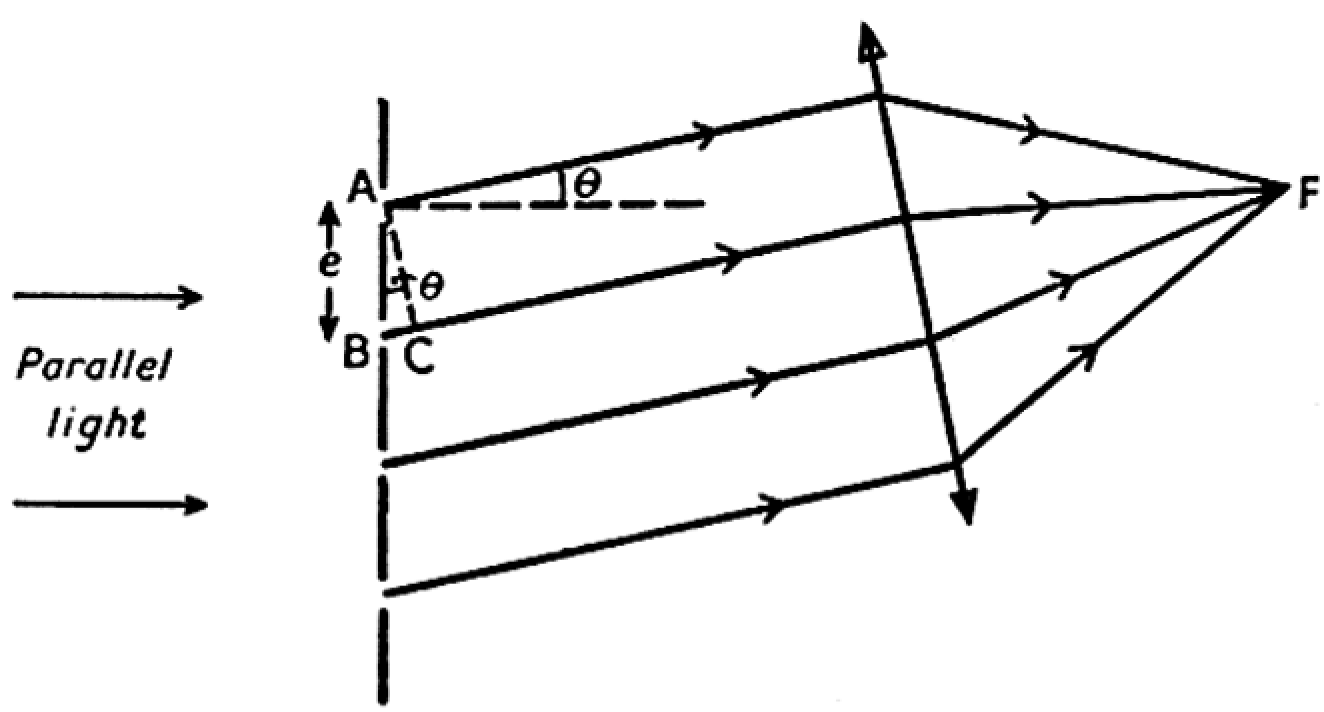


FIG. 734

grating are perpendicular to the paper and, since they are effectively opaque, they are indicated by short black lines. The intervening spaces allow the passage of light and are equivalent to slits. Let e be the distance from a point A in a given position in one slit to the point B in the same position in the next slit. This distance is called the **grating element,** and is equal to the combined width of a line and a space.

We shall confine ourselves for simplicity to the case in which parallel light is incident in a direction at right angles to the plane of the grating, as indicated by the arrows on the left-hand side of the drawing. The light can be produced by an illuminated slit placed at the principal focus of a converging lens, which then sends parallel light to the grating. A lens

arranged on the other side of the grating as shown will bring to its principal focus F all the light which leaves the grating at an inclination θ to the direction of the incident light. The path difference in this direction between light given out as secondary wavelets from A and from B is equal to BC if AC is drawn perpendicular to the direction considered. Therefore if BC is a whole number of wave-lengths, then the wavelets from A and B arrive at F in phase and reinforce each other there. Furthermore, the wavelet from any point in any one slit reinforces that coming from the corresponding points in all the other slits so that there is complete reinforcement. Evidently

$$\begin{aligned} BC &= AB.\sin\theta \\ &= e\sin\theta \end{aligned}$$

so that there is maximum light at F for values of θ which satisfy the equation

$$e\sin\theta = m\lambda$$

where m can be 0, 1, 2 . . . When $m=0$, $\theta=0$, and we have in this case the direct image of the slit formed by the light passing straight through the grating and being brought to a focus at F, the axis of the lens being normal to the place of the grating. This image is flanked on each side by the diffraction images for which m is 1, 2, etc. We refer to these as "images" because in fact they are very sharp maxima. This is due to the fact that each is produced by wavelets coming from the very large number of slits, with the result that when these are all in phase the illumination is comparatively large and falls off rapidly with a small change of θ.

For a grating having 14,000 lines to the inch there are about 5500 lines to the centimetre. Suppose that this is the exact number. Then e is equal to $\frac{1}{5500}$ cm. The value of θ for the first order diffraction is then given by

$$\frac{\sin\theta}{5500} = \lambda$$

Sodium light consists of two nearly equal wave-lengths, namely $5{\cdot}890 \times 10^{-5}$ cm. and $5{\cdot}896 \times 10^{-5}$ cm. For the former, θ comes to be 18° 54′ and for the latter 18° 55′. For the second order diffraction ($m=2$) the corresponding angles are 40° 23′ and 40° 25′, and for the third order 76° 23′ and 76° 37′. If we attempt to work out the angle for the fourth order we find that its sine exceeds unity so that this order does not exist.

Measurements with a diffraction grating are made with the help of a spectrometer. The grating is placed on the spectrometer table so that its lines are parallel to the collimator slit and its plane is perpendicular to the direction of the parallel light coming from the collimator. If the telescope is focused for parallel light, the various diffracted images of the slit will be seen in it in turn as it is swung round. When the telescope

is placed so that its axis coincides with that of the collimator, the direct image of the slit due to light passing straight through the grating is formed on the crosswire as it would be if the grating were absent. This is the maximum for which $m=0$, and consists of a single image whether monochromatic or white light is used because its position does not depend on λ.

By turning the telescope to left or right of the direct position, the first, second and higher orders (if they exist) will be seen in turn. When monochromatic light is used, the diffracted images are comparatively narrow lines, and the angle of diffraction (θ) for any one of them can be measured by noting the reading of the telescope vernier when the particular image is on the crosswire and subtracting it from the reading when the direct image is on the crosswire. The value of θ for any particular order should, of course, be the same on both sides of the direct image. When θ has been determined, λ may be calculated from the equation $e \sin \theta = m\lambda$.

When the slit is illuminated by white light, a full spectrum is obtained in each order instead of a single maximum. According to the equation, the violet end of the spectrum has a smaller angle θ than the longer wave-length red. Consequently the grating spectrum is the other way round from that formed by a prism in which the violet suffers the greatest deviation. The two spectra also differ in respect of the relative widths of particular colour ranges. Since

$$\sin \theta = \frac{m\lambda}{e}$$

differentiating with respect to λ gives

$$\cos \theta \cdot \frac{d\theta}{d\lambda} = \frac{m}{e}$$

or

$$\frac{d\theta}{d\lambda} = \frac{m}{e \cos \theta}$$

The quantity $\frac{d\theta}{d\lambda}$, which is the change of θ per unit change of λ, is the **dispersive power** of the grating, and is practically constant over the whole width of a spectrum of given order m because $\cos \theta$ changes by only a small amount over this range. In other words, equal distances along the spectrum as seen in the telescope represent practically equal changes of wave-length. This spectrum is called a "normal" spectrum. The dispersive power of glass is greater at the violet end than at the red end of the spectrum, so that a spectrum formed by a prism is stretched out at the violet end. It should be noticed that the dispersive power of the grating is proportional to the order m and inversely proportional to the grating element e.

Another point which we can easily deduce is that, since λ for red is approximately double λ for violet, the value of $\sin\theta$ for the red end of the first order spectrum is approximately the same as for the violet end of the second order spectrum. The first and second order spectra are therefore contiguous. It is easily shown by a similar argument that the violet of the third order spectrum begins at the yellow of the second order, so that this overlapping renders observation of the white-light spectrum useless in the higher orders.

Even when the slit is illuminated with monochromatic light the "image" formed by a grating spectroscope is not a sharp image. It is of the same nature as the central maximum of the diffraction pattern due to a single slit (Fig. 733). Therefore it is possible for the difference between two wave-lengths incident on the grating to be so small that the maxima overlap each other and are indistinguishable from each other. In order to make them distinguishable each must be made narrower without decreasing the distance between them. It can be shown theoretically that this can be done by keeping the value of e constant (this maintains the same dispersive power and therefore the same separation of the maxima) and increasing the number of rulings on the grating, *i.e.* increasing the actual width or aperture of the grating. The latter step makes the maxima sharper. Thus there is the same problem of resolving power with the grating as with other optical instruments, such as the telescope and microscope, and the improvement of resolving power is carried out in the same kind of way.

Lastly, we may mention that the diffraction pattern due to each individual slit of a grating is, so to speak, superimposed on the diffraction due to the grating as a whole. Use is made of this fact, for example, in reducing the intensity of one of the orders due to the grating by making the relative widths of the slits and opaque lines such as to cause a minimum of the slit pattern in the direction of the order of the grating pattern which is to be suppressed. More light is concentrated into the other orders in this way.

EXAMPLES LII

1. Explain the production of a pure spectrum by means of a diffraction grating. How is a diffraction grating produced? Find the angular separation of the two sodium lines (wave-lengths $5{\cdot}890$ and $5{\cdot}896 \times 10^{-5}$ cm.) in the first order spectrum produced by a grating with 5000 lines to the centimetre. (C.H.S.)

2. What is meant, in optics, by (*a*) interference, (*b*) diffraction? What part do each of these phenomena play in the production of spectra by a diffraction grating?
A parallel beam of sodium light is incident normally on a diffraction grating. The angle between the two first order spectra on either side of the normal is 27° 42′. Assuming that the wave-length of the light is 5893×10^{-8} cm., find the number of rulings per cm. on the grating. (J.M.B.H.S.)

Chapter LIII

THE SPECTRUM

1. THE SPECTROSCOPE

The Dispersion of White Light by a Prism.—In Vol. 2 of this book (page 527) we have explained that when a single ray of radiation containing different wave-lengths passes through a prism of a material such as glass, the emergent beam consists of a collection of diverging rays, one for each wave-length, as indicated in Fig. 735. Each wave-length arrives at a different point on a screen placed so as to intercept the beam, thus producing a spectrum. It should be noticed that the separation of the rays takes place at the two surfaces where refraction occurs.

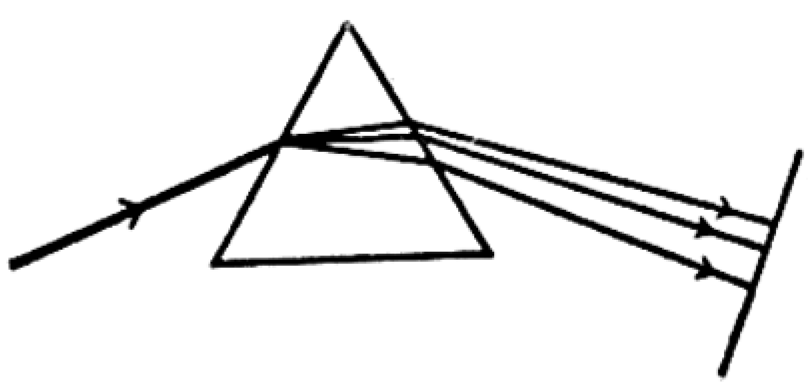


FIG. 735

In practice, of course, it is impossible to confine the incident radiation to a single ray. Without using any focusing device, such as a lens, the nearest approach to a single ray would be a narrow beam such as would issue from a pinhole (or slit parallel to the refracting edge of the prism) placed at a considerable distance from a source of light. Newton, in his famous experiments on the spectrum, used a small hole in a window shutter to produce a narrow beam of sunlight. The screen on which the light fell after passing through his prism was at a distance of about 18 ft. from the prism, and in the absence of the prism an image of the sun, about $2\frac{1}{8}$ in. in diameter, was formed on the screen after the fashion of a pinhole camera. In Newton's experiment, therefore, *each* wave-length in the incident radiation produced a *patch* of light in his spectrum (which was about 10 in. long), so that the patches due to wave-lengths differing by anything less than a certain considerable amount overlapped. Since there is a continuous gradation of wave-length in sunlight there is much overlapping in such a spectrum, and any one point on the screen receives not one wave-length but a range of wave-lengths.

Nevertheless the band of light on Newton's screen showed a continuous gradation of colour from violet (the most deviated) to red, and Newton concluded that white sunlight was a mixture of the colours in his spectrum, and that their separation was due to the fact that the refractive index of the glass of the prism was different for each colour. In an experiment to verify this, Newton made a small hole (H_1, Fig. 736) in a screen on which

a spectrum was being formed by a prism P_1. A restricted range of wave-lengths passed through this hole and then through a second hole H_2 in another screen. The light from H_2 was passed through a second prism P_2 and finally fell on a screen S. The holes H_1, H_2 and the prism P_2 were held in fixed positions, so that the direction of the light incident on P_2 was constant. If P_1 was then rotated so that the spectrum moved up the first screen with the result that the colour of the light entering H_1 moved towards the violet end of the spectrum, it was observed that the patch of light on S moved downwards, thus showing that the deviation due to P_2 was increasing as the colour of the light approached violet. This confirms the fact that the formation of the original spectrum by P_1 is due to the variation of the deviation with colour.

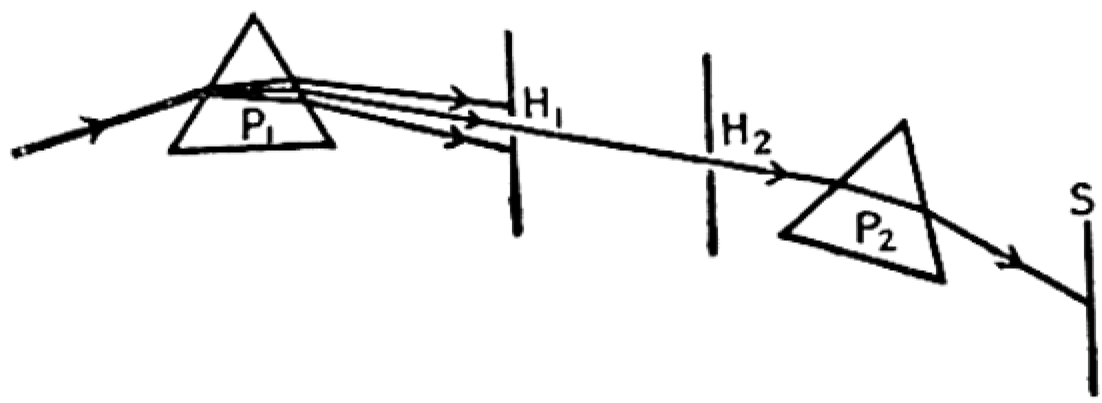


FIG. 736

Another experiment Newton did to illustrate the same thing is shown in Fig. 737. The prism P_1 produced the usual dispersion of the incident white light, which then passed into a second prism P_2 whose refracting edge was perpendicular to that of P_1. The deviation due to P_2 was therefore at right angles to that produced by P_1 and, as indicated in the figure, the spectrum was formed slantwise on the screen S because the deviation due to P_2 increased from the red to the violet end of the spectrum.

It is possible, of course, to recombine the spectral colours into white light. One way of doing this has already been referred to on page 768. The dispersed white light is allowed to fall on a second prism identical with the first. The refracting edges of the two are parallel, but the second prism is inverted with respect to the first. As already mentioned (page 769), both the dispersion and the deviation are neutralized by the second prism when the two are identical, but, by using two types of glass with different dispersive powers and choosing the right values of the refracting angles, either dispersion or deviation can be eliminated. The reader's attention is called to the discussion on page 768 of the character of the emergent beam. The spectral colours may also be recombined by a converging lens (page 980).

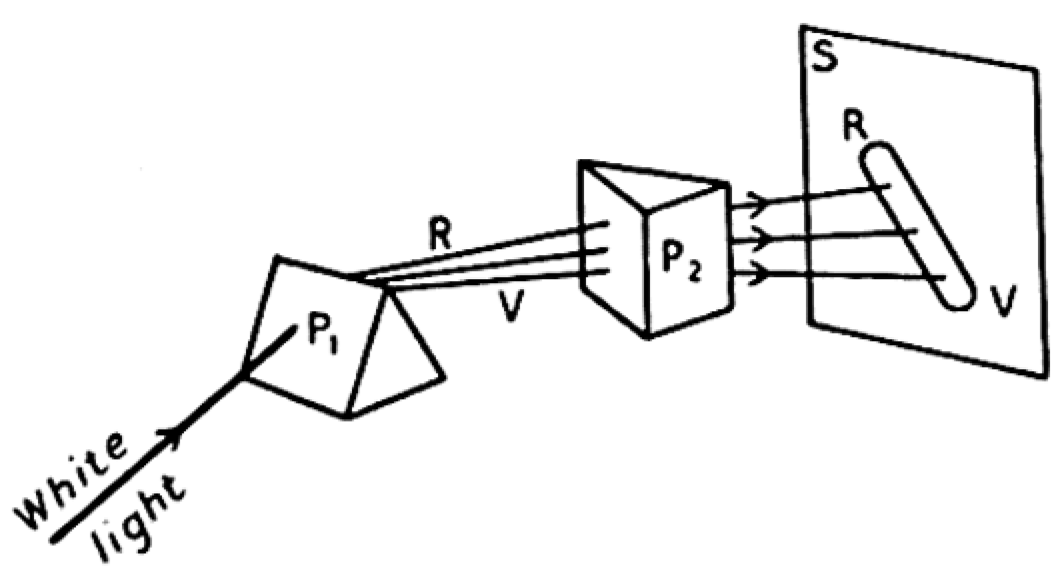


FIG. 737

The Focused Spectrum.—It has already been pointed out that there is overlapping when each colour present in the white light incident on a

prism would, by itself, create a patch of finite size on the screen. The light arriving at any one point in the spectrum is not monochromatic but contains a range of wave-lengths. Steps are therefore taken to reduce this range as much as possible. If it could be reduced indefinitely so that every point in the spectrum represented a single wave-length, the spectrum would be described as **pure.**

To increase the purity of a spectrum it is obviously necessary to reduce the size of the patch of light which each wave-length produces on the screen, and this means reducing the size of the white-light patch which would be produced on the screen in the absence of the prism. This can be done by placing a narrow slit between the lamp (or sun, etc.) and the prism, the slit being parallel to the edge of the prism, and then inserting a converging lens to form a real image of the slit on the screen. The arrangement is shown in Fig. 738. In the absence of the prism the lens

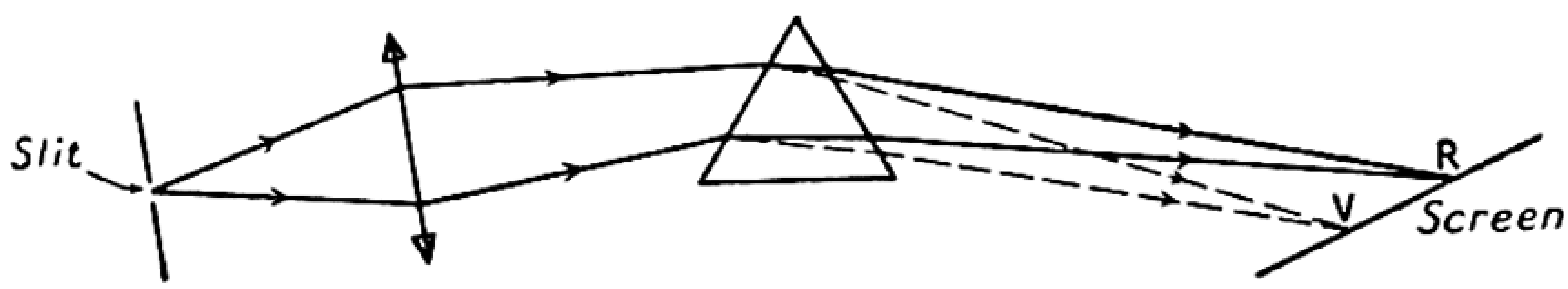


FIG. 738

would, of course, form a real white-light image of the slit. When the prism is present, each wave-length in the incident light produces an image of the slit and the spectrum consists of these images. Because of the necessarily finite width of the slit, and therefore of its image, the images due to wave-lengths which differ by less than a certain amount are still bound to overlap. Nevertheless, because of the narrowness and sharpness of the slit images compared with the patch of light such as is obtained by Newton's original experiment, the spectrum is correspondingly purer. It is called a "focused" spectrum because it consists of a series of focused images.

In fact, however, the focusing is by no means perfect owing to the fact (which is obvious from Fig. 738) that the rays reaching the prism from the lens do not all strike the prism at the same angle. The graph of deviation (δ) against angle of incidence (i) for any particular colour in the incident light is of the type shown in Fig. 739 (i). It shows the characteristic minimum deviation (page 753) which occurs at a particular value of i. Fig. 739 (ii) shows three cases of narrow converging pencils passing through a prism. Suppose that the convergent rays labelled (1) and (2) have, respectively, angles of incidence i_1 and i_2 (shown in Fig. 739 (i)) which are respectively slightly less and slightly greater than the angle of incidence for minimum deviation. According to the graph, the deviation of each of these rays is the same, so that their inclination to each other is the same on leaving the prism as on entering it. Thus

the convergence of the narrow pencil bounded by rays (1) and (2) is unaltered by its passage through the prism. The pencil itself can be regarded as suffering the minimum deviation. Another pair of rays, (3) and (4), have angles of incidence less than that required for minimum deviation. Each suffers a deviation greater than the minimum, but by marking their angles of incidence i_3 and i_4 on the graph, we see that ray (3) suffers a greater deviation than ray (4), with the result that the convergence of the pencil is increased and it comes to a focus at I_3, which is nearer the prism than I_1. Similarly for the pair of rays (5) and (6) the angles of incidence are i_5 and i_6 and the deviation of (6) is greater than that of (5), so that their focus, I_5, is further from the prism than I_1.

When a similar argument to the above is applied to *diverging* pencils it is found that, like the convergence of a converging pencil, the divergence of a diverging pencil is decreased by passing through the prism when the angle of incidence is greater than that for minimum deviation. For an angle of incidence smaller than that for minimum deviation the divergence is increased. In general, therefore, we can say that a pencil of rays (whether diverging or converging) which is incident at an angle greater than that for minimum deviation is rendered more nearly parallel, while for an angle of incidence smaller than that which gives minimum deviation the pencil is made less nearly parallel.

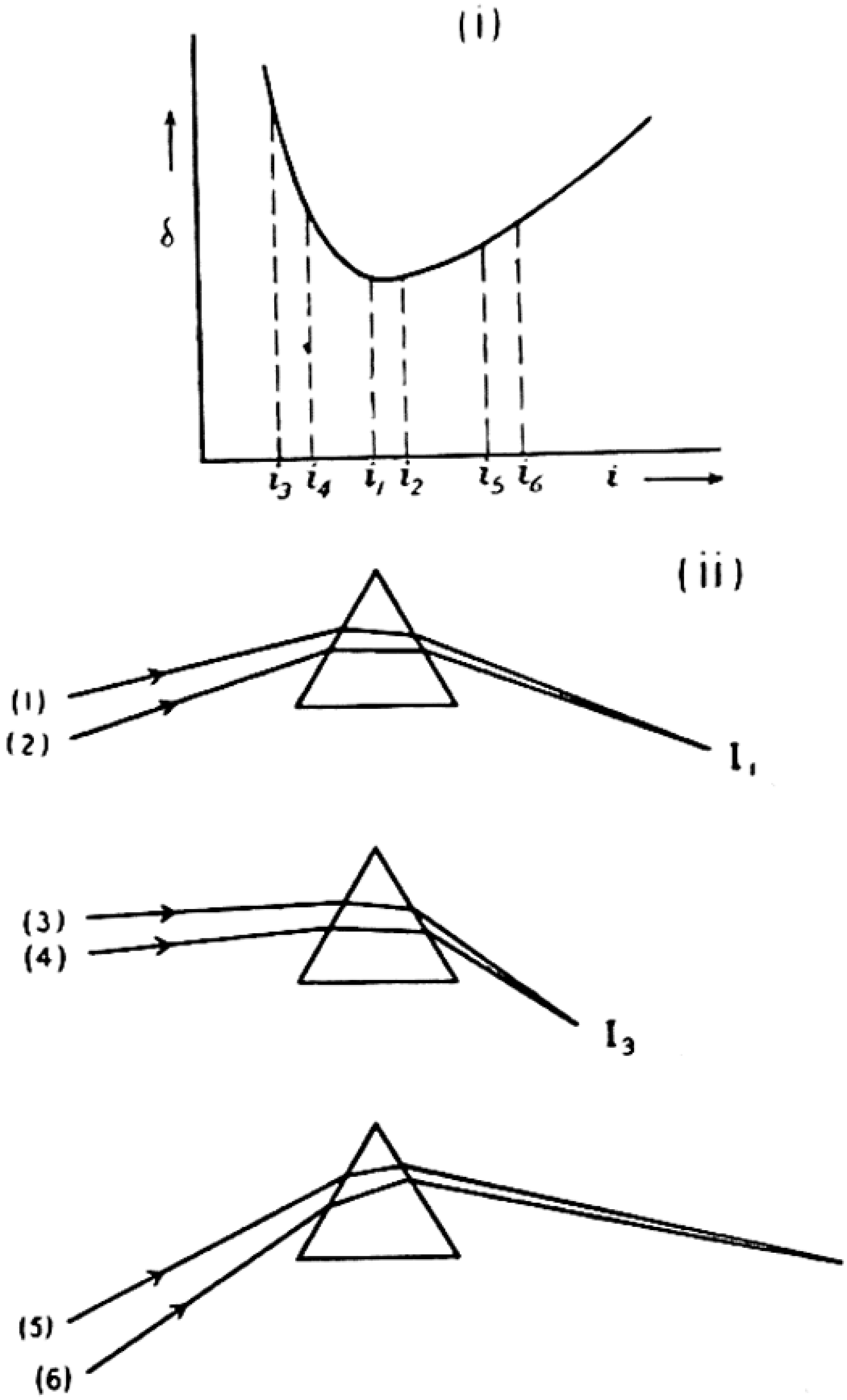


Fig. 739

Thus, if we imagine that the fairly wide beam of a given colour striking the prism in Fig. 738 is divided into small pencils of varying angles of incidence, we see that the separate images of the slit produced by the pencils do not coincide with each other. The focusing is therefore imperfect, although this is a small effect compared with the variation of deviation with colour, so that a fairly satisfactory spectrum is produced.

The student may verify for himself that since violet is the most refrangible light in the spectrum, the angle of incidence for minimum deviation of violet by a given prism is greater than for any other colour. Therefore if a convergent pencil of white light (Fig. 738) is incident on a prism at an angle for which the deviation of the centre of the spectrum (say yellow) is a minimum, the violet light in it is incident at an angle which is less than that required for minimum deviation of this colour. It is therefore made more convergent, and comes to a focus nearer the prism than the other colours of the spectrum. Thus the best focusing of the spectrum is obtained by sloping the screen as shown in Fig. 738.

The imperfect focusing of convergent or divergent beams incident on a prism can be eliminated by using an incident beam of *parallel* rays. Each ray of a given colour then strikes the prism at the same angle of incidence, so that each suffers the same deviation, and they emerge from the prism parallel to each other. They can then be brought to a sharp focus in the focal plane of a converging lens. The arrangement is shown in Fig. 740.

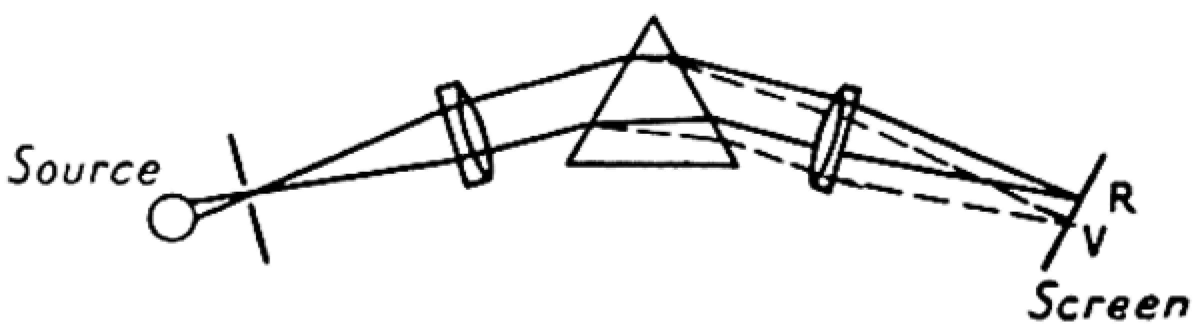


FIG. 740

The first lens renders parallel the white light from any one point on the slit, and the second lens brings to a separate focus on the screen the parallel rays of each colour emerging from the prism. The spectrum is a series of images of the slit, and its purity is increased when overlapping is diminished by narrowing the slit. It can never be perfectly pure because the slit, and therefore its image formed by one particular colour, cannot be made indefinitely narrow. It should be noticed that the lenses are represented in the diagram as achromatic doublets. Lenses uncorrected for chromatic aberration would introduce unwanted dispersion.

The Spectroscope or Spectrometer.—The arrangement shown in Fig. 740 is the basis of the prism spectroscope. A form of spectroscope or spectrometer (the latter name is applied when the instrument is used to determine the refractive index of the material of the prism rather than to examine spectra) is shown diagrammatically in plan in Fig. 741. The **collimator** consists of a pair of telescopic tubes with a slit of adjustable width mounted at one end and an achromatic converging lens at the other. This arrangement gives parallel light from each point on the slit when the distance between the slit and the lens is adjusted to be equal to the focal length of the latter. The light from the collimator falls on the prism, which is placed on a turn-table T which can be levelled by means of three levelling screws. The vertical axis about which the turn-table

rotates should intersect the axis of the collimator, and the angle through which the table turns relative to the base of the instrument can be measured by means of a vernier attached to it which moves over a circular scale graduated in degrees and submultiples.

The light which is refracted and dispersed by the prism can be received by a telescope having an achromatic objective, in the focal plane of which a real image of the spectrum is formed. The eyepiece of the telescope forms a virtual magnified image of the spectrum. The telescope can be focused by altering the distance between the objective and eyepiece, and it has a pair of crosswires which, when the instrument is properly adjusted, are in the plane of the real image of the spectrum. The telescope can be rotated, as indicated by the arrows in the drawing, about the axis of the

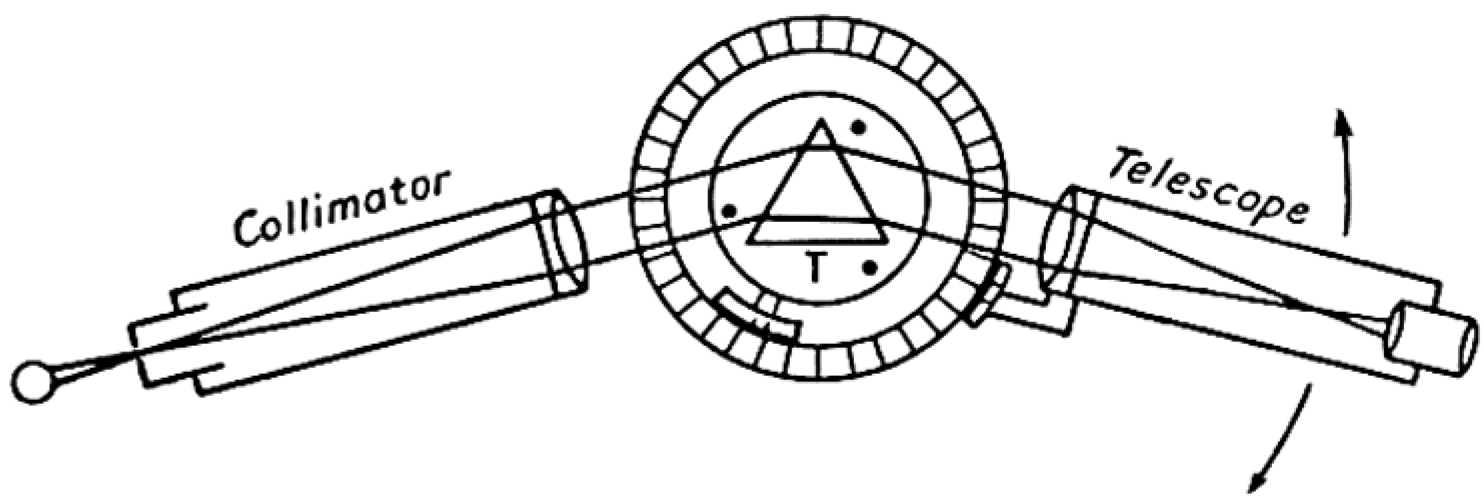


FIG. 741

turn-table, and the angle through which the axis of the telescope turns can be measured on the circular scale by means of the vernier which is attached to the telescope. In many instruments two such verniers separated by 180° are used for greater accuracy, both on the telescope and on the turn-table.

We shall now describe the adjustment and use of the spectrometer for the determination of the refractive index of the material of a prism. The details of the experiments can be satisfactorily learnt only by their actual performance, so that we shall confine ourselves to the explanation of the principles.

Adjustment of the Spectrometer.—In adjusting a spectrometer it is necessary to use a monochromatic light source, *e.g.* light from a sodium flame, placed on the axis of the collimator and at such a distance from the slit that the jaws of the latter are not overheated. The slit may be made quite wide at first, and if, before placing the prism on its table, the telescope is turned so as to receive light direct from the collimator, a patch of light will be seen. The crosswires are viewed against this background, and the eyepiece is pushed in and out of the telescope until they are seen clearly in focus. This adjustment is peculiar to each individual observer, since it is affected by whether his sight is normal or not, and whether he prefers to have the image of the crosswires at his near point or far point. The latter is the more restful, of course, because it does not involve accommodation of the eye.

It is necessary to focus both the telescope and collimator for parallel light, and an easily understood way of doing this is as follows. The telescope is pointed at the most distant object available, and the distance from the objective to the crosswires is adjusted until the object is seen clearly and, finally, until there is no parallax between the image of the object and that of the crosswires as the observer moves his head from side to side. Since the object is very distant, it forms a real image in the focal plane of the objective, and since the crosswires have been made to coincide with this image, the adjustments carried out so far have ensured that when the eye sees the crosswires clearly through the eyepiece, it is also receiving light which was parallel before entering the telescope. The telescope is then turned to receive the direct light from the collimator, and the latter is focused until the image of the slit is clearly seen in the telescope when the eye is looking at the crosswires and, for greater accuracy, until there appears to be no parallax between the image of the slit and that of the crosswires. The slit should be made narrow during the final focusing.

An alternative method of focusing the telescope and collimator for parallel light, known as **Schuster's method,** requires the prism to be in place on its table. Light from the collimator is made to fall at a suitable angle on the first face of the prism, and the telescope is turned so as to receive the deviated light. An eye looking into the telescope then sees an image of the slit which, in general, will be blurred because neither the collimator nor the telescope is focused. The prism is then turned and the telescope is moved so as to follow the blurred image until the position of minimum deviation is reached. This is recognized by a reversal of the direction of movement of the image while the prism is being turned continuously in one direction. The telescope is then so placed that the minimum deviation position of the image is a little to one side of the centre of the crosswires, and the image is on the crosswires when the deviation is a little more than the minimum, which will occur for *two* positions of the prism, *i.e.* for two different angles of incidence of the light striking the prism from the collimator.

The adjustment now proceeds by focusing the *telescope* so as to give a clear image of the slit when the angle of incidence is *greater* than that for minimum deviation (call this position X of the prism), and the prism is then turned so that the angle of incidence is *less* than that for minimum deviation (position Y). In general, the slit image will no longer be in focus. On page 961 it was shown that in position X the prism causes the rays passing through it to become more parallel on emergence and in position Y less parallel. Therefore in position Y the *collimator* alone is focused so as to give a clear slit image in the telescope. This restores the convergence or divergence of the rays leaving the prism to what it was in position X (since the telescope was focused in this position) and causes the rays to be more nearly parallel as they leave the collimator. We now return to position X. The rays leaving the prism are now made still

more parallel by their passage through it, and an adjustment of the *telescope* alone for a sharp image now brings this part of the instrument nearer to the condition in which it is focused for parallel light. Again the prism is turned to position Y and the collimator is focused, then back to position X for a further adjustment of the telescope, and so on until the slit is in clear focus in both positions without further adjustment. The light leaving the collimator must then be parallel, and the telescope must therefore be focused to receive parallel light.

The advantage of Schuster's method of focusing a spectrometer lies in the fact that it is not necessary to use a distant object, which is often inconvenient if the spectrometer is not easily moved. It is not necessary to remember which of the two, telescope or collimator, has to be adjusted in any one position of the prism. The focus of the slit image rapidly gets worse if the process is carried out in the wrong way.

It should be mentioned that in order to achieve the sharpness of focus which is necessary in a spectrometer, the surfaces of the prism must be really plane, otherwise it is impossible to obtain a sharp slit image by refraction through or reflection from them. The glass of the prism must also be homogeneous for the same reason. Prisms which are good enough for experiments with pins may be quite unsatisfactory in a spectrometer.

The last adjustment of the spectrometer is known as "levelling the prism." This means, in effect, making the planes of the refracting surfaces of the prism perpendicular to the plane containing the axes of the collimator and telescope. This can be done by adjusting the screws attached to the prism table until the refracted slit image does not move up or down in the field of view as the prism table is turned and the image is followed in the telescope. In good spectrometers the directions of the axes of the collimator and telescope can be adjusted in a *vertical* plane by means of screws, and a spirit level is sometimes used to make the axes horizontal. It is not possible, however, for us to go into a detailed account of all the possible adjustments which may have to be made in accurate work with expensive instruments.

Use of the Spectrometer. *Measurement of Refracting Angle.*—The determination of the refractive index of the material of a prism from measurements of minimum deviation and refracting angle made by ray-tracing with pins has already been described (page 755). The measurement of the angles can, of course, be made much more accurately with a spectrometer, and it is necessary to use monochromatic (*e.g.* sodium) light. The principle is, however, the same as that described on pages 752–756.

The refracting angle α of the prism may be measured by placing the prism on the table of a properly adjusted spectrometer with its refracting edge towards the collimator as shown in Fig. 742. This causes the parallel rays to be divided into two reflected beams. The telescope is placed so as to receive one of these, and turned until the reflected image of the slit is on the centre of the crosswires, say position T_1 of the telescope

in Fig. 742. The reading of the telescope vernier is taken, and the telescope is then turned to the position T_2 in which the image of the slit formed by reflection by the face AC is on the crosswires. The vernier is again read and the difference of the readings gives the angle θ. As explained on page 756, the refracting angle of the prism a is given by

$$a = \frac{\theta}{2}$$

Another method consists in keeping the telescope fixed and noting the angle through which the prism must be turned in order to bring on to

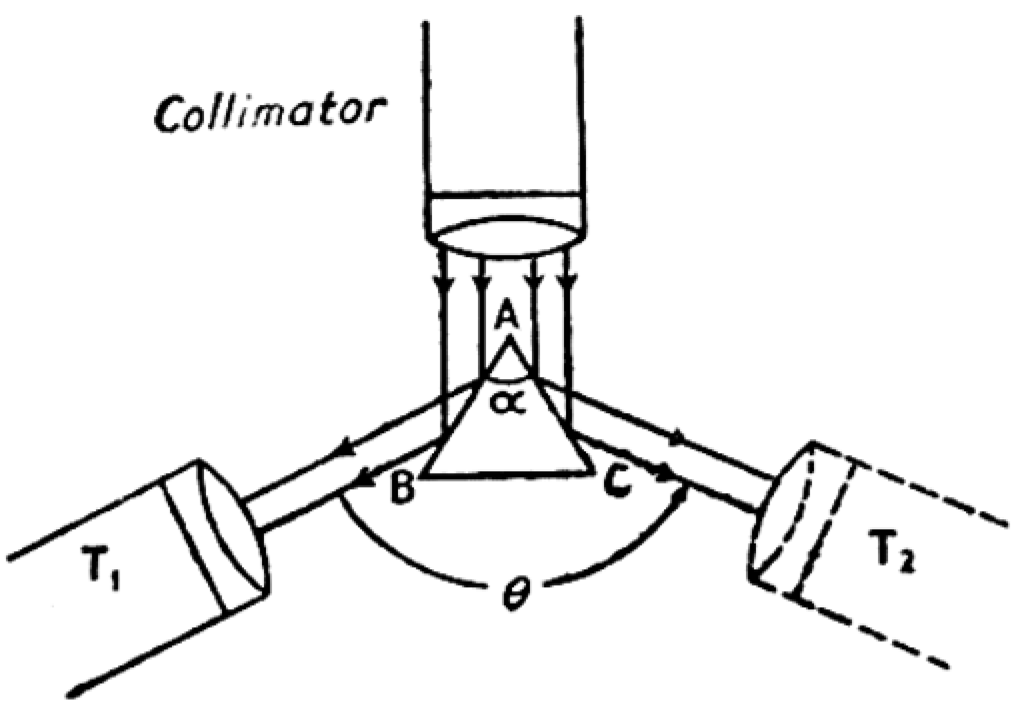


FIG. 742

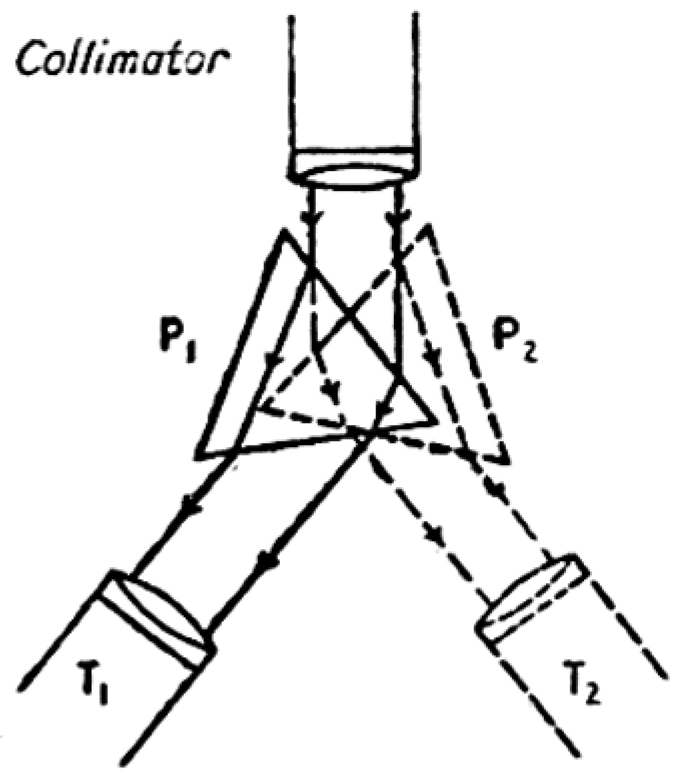


FIG. 743

the crosswire first the image formed by reflection from AB and then from AC. The value of a is the supplement of the angle of rotation of the prism.

Measurement of Minimum Deviation.—The angle of minimum deviation can be found by first noting the reading of the telescope vernier when, in the absence of the prism, the telescope is placed so as to receive the direct image from the collimator on the crosswires and, second, taking the reading when the telescope has been turned to receive the minimum-deviation image on the crosswires. The difference of these readings is the minimum deviation. It is more accurate, however, to take the two readings for minimum deviation on opposite sides of the direct position. In going from one to the other the prism is turned from position P_1 to P_2 (Fig. 743), and the telescope goes from T_1 to T_2. The difference of the readings of the telescope vernier for T_1 and T_2 is equal to twice the angle of minimum deviation.

When the angle of the prism and minimum deviation have been determined, the refractive index of the material of the prism for the particular light used can be calculated from the usual formula (page 755). The refractive index of a liquid can be found by placing a quantity of it in a **hollow prism** (page 760).

When a gas contained at low pressure in a glass tube is subjected to a high-voltage electrical discharge it is made luminous, and if such a tube containing, say, helium is placed in front of a spectrometer slit a **line spectrum** is observed. This is because the gas emits a number of separate wave-lengths, each of which, on account of the dispersion in the prism, produces a separate image of the slit. Each line occupies the same position as would light of the same wave-length (colour) in the continuous spectrum of a white-light source. Thus we can measure the minimum deviation for each of the lines and so obtain the refractive index of the material of the prism for each. Taking a pair of extreme lines, one red and one blue or violet, and a central line (yellow), we can calculate the dispersive power of the material of the prism (page 767). It is given by

$$\frac{(\text{refractive index for violet}) - (\text{refractive index for red})}{(\text{refractive index for yellow}) - 1}$$

Furthermore, a graph can be plotted of the refractive index of the prism material against the wave-length of the light *in vacuo*, since the wave-lengths of the lines in the spectra of various substances are known. For glass a curve such as Fig. 744 is obtained, and the results can be expressed by an approximate form of Cauchy's equation

$$n = A + \frac{B}{\lambda^2}$$

in which A and B are constants for any given material. If the equation is true, a graph of n against $\frac{1}{\lambda^2}$ would be a straight line, and, to within the degree of experimental accuracy obtainable with an ordinary laboratory spectrometer, this will be found to be the case over all but the extreme regions of the visible spectrum. The full Cauchy equation includes a third term $\left(\frac{C}{\lambda^4}\right)$ on the right-hand side.

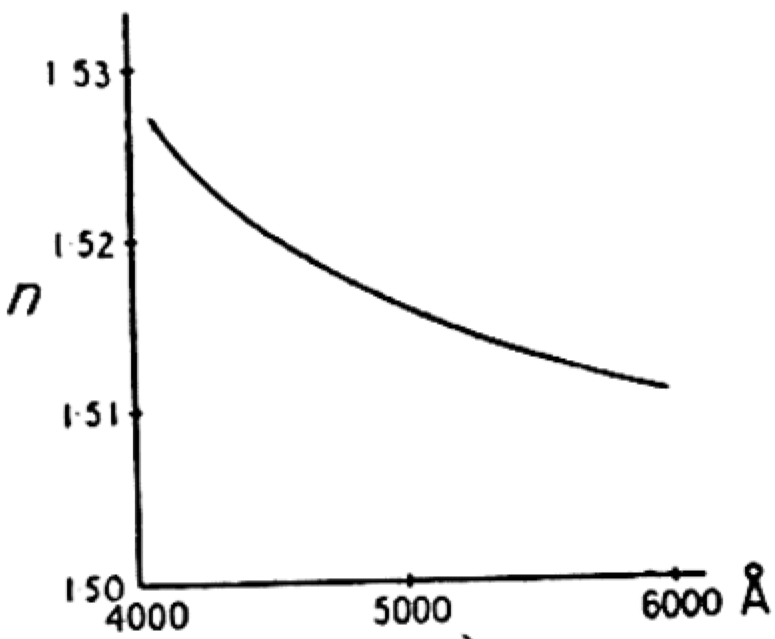


FIG. 744

Examination of Spectra.—When the instrument is used primarily for examining spectra rather than determining refractive index, it is more suitably called a spectroscope. A spectrograph is a spectroscope adapted for photographing spectra. This can be done by removing the telescope eyepiece and substituting the lens of a camera, the plate of which will receive an image of the first real image of the spectrum.

Spectroscopy is concerned with the examination of the nature of spectra and with the measurement of the wave-lengths of spectral lines. Absolute wave-length measurements cannot be made with a prism spectroscope, but a diffraction grating may be substituted for the prism, in which

case, if the number of lines per centimetre on the grating is known, the wave-length of any line in a spectrum can be calculated (page 955) from the angle of diffraction of the particular line. This angle can be found by arranging the plane of the grating perpendicular to the axis of the collimator and moving the telescope until the diffracted image of the slit (*i.e.* the spectral line) of a given order is on the crosswires. The angle of rotation of the telescope from this position to the corresponding position on the other side of the direct image is noted, and the angle of diffraction for the given order is then half the angle of rotation of the telescope. Thus if the wave-lengths of a sufficient number of spectral lines are found by this or other methods they can be used to calibrate a prism spectroscope. A convenient light source for this calibration is a tube containing helium at low pressure through which an electrical discharge is passing. This gives six or eight bright, easily recognizable lines distributed throughout the spectrum. One way of effecting the calibration consists in setting the prism in the minimum-deviation position for, say, the yellow line and noting the deviations (which will not be minimum) for the other lines and plotting a graph of deviation against wave-length. This can subsequently be used to determine the wave-length of an unknown line, provided that the same prism is used in the same position. Alternatively, minimum deviation may be measured for each line and plotted against wave-length, or a graph may be constructed of refractive index against wave-length. In order to obtain a straight-line graph, and therefore to assist interpolation, refractive index may be plotted against the reciprocal of the square of the wave-length (Cauchy's formula).

2. TYPES OF SPECTRA

Spectra are described by various terms, some of which apply to the method of excitation (*i.e.* the method of rendering luminous) of the source of the radiation, while others are descriptive of the appearance of the spectra. All spectra which represent the analysis of the light emitted by a luminous body are known as **emission spectra** to distinguish them from absorption spectra which we shall describe later on.

The Continuous Spectrum.—A continuous spectrum is one in which there are no gaps and no abrupt changes of intensity. Spectra of this type are given by heated solids, *e.g.* the filament of an electric lamp, the crater of a carbon arc, etc. The distribution of energy in the spectrum of an ideal black body and its variation with temperature have already been discussed (page 528, Vol. 2). Curves of a similar shape to those shown in Fig. 391 are obtained with the radiation from actual incandescent sources even although they are not perfectly black. The quality of the radiation given out by a heated solid depends more on its temperature than on the nature of the body itself, and the radiation is regarded as being produced by the vibrations of the individual molecules of the

substance which, of course, are made more rapid when the temperature is raised, thereby increasing the proportion of high-frequency (short wave-length) waves emitted.

Methods of Exciting Luminosity.—We shall now briefly mention some of the ways in which substances are made luminous for the purpose of examining the spectra which they give.

Incandescent Vapour.—The commonest example of the use of an incandescent vapour in spectroscopy is the sodium flame, to which reference has often been made. In this, a sodium salt is placed in the base of a non-luminous bunsen flame. The flame emits the familiar yellow light characteristic of sodium and, when it is used to illuminate the slit of a spectroscope, the spectrum is seen to consist of two sharp yellow lines so close to each other (*i.e.* of so nearly equal wave-length) that a small spectroscope often fails to separate them. For many purposes, therefore, sodium light can be regarded as monochromatic. Salts of other metals, *e.g.* lithium, strontium, potassium, etc., give light of characteristic colours when placed in a flame. This fact can be used to identify the metals, and the presence of any one metal can be detected with certainty by using a spectroscope to measure the wave-lengths of the lines in the spectrum of the flame and comparing them with tables of wave-lengths which are known for all metals.

Metal vapours can also be made incandescent by striking an electric arc between electrodes of the metal or by placing some of the metal in the crater of a carbon arc. In both cases it is the flame of the arc which emits the **arc spectrum** of the metal.

Discharge Tube.—In this piece of apparatus, a gas at low pressure (a few millimetres of mercury) is contained in a tube such as that shown in Fig. 745 with metal electrodes sealed into its ends. An electrical potential difference is applied to the electrodes, and a glow is then produced in the tube which has a colour characteristic of the gas. The narrow central part of the discharge tube (where the light is brightest) is used as a light source for a spectroscope, the tube being placed parallel to the slit.

FIG. 745

Spark Spectra.—A spark spectrum is the spectrum produced by the light of an electric spark discharge passing between electrodes of the particular metal.

Line and Band Spectra.—A line spectrum is a collection of separate sharp lines. Each line is an image of the slit formed by light of a particular wave-length, thus showing that the light entering the spectroscope is a mixture of these separate distinct wave-lengths. As already stated, line spectra are given by metallic vapours made incandescent in a flame. The arc and spark spectra of metals are also line spectra, the latter containing many additional lines not seen in the former.

The light from an electric discharge through rarefied gases often gives a line spectrum. The simplest of these is the hydrogen spectrum which, in the visible range, consists of one red line (the C line, which has a wave-length 6562·8 Å), one blue-green line (F, 4861·3 Å), and two violet lines (4340·5 Å and 4101·7 Å). Helium gives more visible lines than hydrogen, and a small helium discharge lamp, capable of being operated by the usual mains voltage, is very useful in calibrating a prism spectroscope, the wave-lengths of the lines being tabulated in books of physical constants.

Band spectra get their name from the separate regions or bands, each of which extends over a considerable wave-length range. One edge of each band is diffuse and the other, which is known as the "head" of the band, is sharp. When high-resolution spectroscopic methods are used, it is found that bands are really series of only slightly separated lines which become more and more crowded together towards the head. Band spectra are found to be due to the vibrations of molecules consisting of two or more atoms. An electrical discharge through nitrogen or carbon dioxide gives a band spectrum, and so does a flame in which calcium salts are being heated, the bands being different for different salts.

Spectral Series.—In 1885 **Balmer** discovered that the hydrogen spectrum contained a series of lines whose wave-lengths could be expressed with considerable accuracy by the equation

$$\frac{1}{\lambda}=R\left(\frac{1}{4}-\frac{1}{n^2}\right)$$

where R is a constant, which later came to be called the **Rydberg constant,** and n is a whole number equal to 3 for the visible line with the longest wave-length (the red C line), 4 for the next (the F line), and so on, the series extending beyond the four visible lines to include lines in the ultra-violet which, as shown by the equation, become more and more crowded together towards the short wave-length end of the spectrum. Balmer's was the first of many similar discoveries. As regards the hydrogen spectrum, there is a series of lines in the ultra-violet, distinct from the Balmer series and known as the **Lyman series,** which obeys the equation

$$\frac{1}{\lambda}=R\left(\frac{1}{1}-\frac{1}{n^2}\right)$$

where R is the same constant and $n=2, 3\ldots$ In the infra-red part of the hydrogen spectrum there is the **Paschen series** given by

$$\frac{1}{\lambda}=R\left(\frac{1}{9}-\frac{1}{n^2}\right)$$

where $n=4, 5, 6\ldots$, and there are other series. The separate series

can be incorporated in the single equation

$$\frac{1}{\lambda}=R\left(\frac{1}{n_2{}^2}-\frac{1}{n_1{}^2}\right) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where R always has the same value, $n_2=1$ for the Lyman series (with $n_1=2, 3, 4 \ldots$), $n_2=2$ for the Balmer series (with $n_1=3, 4, 5 \ldots$) and $n_2=3$ for the Paschen series ($n_1=4, 5, 6 \ldots$).

The explanation of these regularities in the spectrum of hydrogen was given in terms of ideas put forward by Rutherford and by Bohr which immediately proved extremely fruitful. Like all hypotheses, however, they have since needed development and modification.

In order to explain the results of experiments on the scattering of α-particles by matter, Rutherford suggested that an atom consisted of a small nucleus in which was concentrated nearly all the mass of the atom and which carried a charge of positive electricity. Around the nucleus and at comparatively large distances from it, electrons (*i.e.* particles of negative electricity) circulated in orbits in the same sort of way as the planets circulate round the sun. The atom as a whole is electrically neutral, so that the total charge of the orbital electrons must be equal to the positive charge on the nucleus. According to Rutherford's model of the atom, the simplest of all possible atoms is that of hydrogen, in which a single electron circulates round a nucleus consisting of a single proton—a particle having a mass about 1840 times that of an electron and a positive charge equal in magnitude to the negative charge of an electron.

Evidently in order to move the electron away from the nucleus, *i.e.* to increase the radius of the orbit, it would be necessary to do work against the mutual attraction of the electron and the nucleus. Therefore the potential energy of the atom is increased if the radius of the orbit is increased.

This very brief description of the Rutherford atom will suffice for our immediate purpose, which is the explanation of spectral series. One apparent difficulty with regard to this model is the fact that on classical theory an accelerated electron would be expected to radiate energy, so that the orbital electrons, being accelerated on account of their curved paths, should radiate continuously at the expense of the potential energy which they have in respect of their attraction towards the positive nucleus. They would therefore approach the nucleus in a spiral path and every atom would, so to speak, collapse. It is here that Bohr's hypothesis enters into the discussion. Bohr applied the ideas of the quantum theory which had already been successful in explaining the facts about thermal radiation (page 543, Vol. 2). Briefly, the quantum theory states that the energy of a system can only have certain values, and that the emission of radiation is not a continuous process but consists in the giving out of small quanta of energy of definite magnitude.

Thus Bohr supposed that the energy of the atom could only have

certain values, which means that the radius of the electron orbit cannot change continuously, but only by jumps from one permitted value to another. The smallest of these permitted orbits represents the lowest energy state of the atom, which is called the **ground state.** When energy is communicated to an atom in its ground state by any of the methods for rendering the element luminous which have already been mentioned, the electron is transferred to a larger orbit of higher energy and the atom is said to be **excited.** In this state the electron is liable to jump to an orbit of lower energy, and the energy which it then loses appears as an equal amount of radiant energy. It was further postulated by Bohr that the frequency ν of the radiation given out during a transition from a higher to a lower orbit was related to the difference ΔE between the energies of the two orbits by the quantum relationship

$$h\nu = \Delta E$$

where h is Planck's constant (page 543, Vol. 2), which has a value of about $6{\cdot}6 \times 10^{-27}$ erg sec. Thus, with the orbits having discrete and definite energies, only certain values of ΔE, and therefore of ν, are possible. This hypothesis therefore explains why a *line* spectrum is produced, and not a continuous spectrum.

Bohr calculated that the energy of an electron in the nth orbit (the smallest orbit being the first) is given by

$$-\frac{2\pi^2 me^4}{h^2}\cdot\frac{1}{n^2}$$

where m is the mass of the electron and e is its electric charge. The negative sign is due to the fact that there is zero energy when the electron is at an infinite distance from the nucleus, so that, since the energy decreases as the distance diminishes, the energy of each orbit must be negative. It follows from the above expression that when the electron falls from an orbit for which $n = n_1$ to an orbit for which $n = n_2$ its decrease of energy ΔE is given by

$$\Delta E = \frac{2\pi^2 me^4}{h^2}\left(\frac{1}{n_2^{\,2}} - \frac{1}{n_1^{\,2}}\right)$$

This amount of energy is therefore emitted as radiation and is equal to $h\nu$, so that the frequency of the radiation is given by

$$\nu = \frac{2\pi^2 me^4}{h^3}\left(\frac{1}{n_2^{\,2}} - \frac{1}{n_1^{\,2}}\right)$$

Since $c = \nu\lambda$, where c is the speed of propagation and λ is the wave-length of the radiation, we have

$$\frac{1}{\lambda} = \frac{2\pi^2 me^4}{ch^3}\left(\frac{1}{n_2^{\,2}} - \frac{1}{n_1^{\,2}}\right) \quad . \quad . \quad . \quad . \quad (2)$$

This equation is identical in form with equation (1), which incorporated all the various spectral series given by hydrogen. The value of the expression $\frac{2\pi^2me^4}{ch^3}$, which in equation (2) replaces the Rydberg constant R of the empirical equation (1), can be calculated from independent values for m, e, c and h, and it is found to be in good agreement with the value of R determined directly from wave-length measurements on the hydrogen spectrum. It will be remembered that equation (1) represents the Lyman series if $n_2=1$ and $n_1=2, 3, 4 \ldots$ Thus in the Rutherford-Bohr model of the hydrogen atom, the first line of the Lyman series is emitted by an atom when the electron falls from the second orbit back to the first (ground state), the second line of the series represents a transition from the third orbit to the first, and so on. Similarly, the various lines in the Balmer series ($n_2=2$, $n_1=3, 4, 5 \ldots$) are produced by electrons falling back from the various possible orbits into the second orbit. The lines of the Paschen series are due to electrons falling back into the third orbit, and so on for all the other possible series. Thus each spectral series is a collection of wave-lengths which are due to the transition of the electron *into* a given orbit from all the possible orbits of higher energy. This is illustrated schematically in Fig. 746, in which the horizontal lines represent the possible **energy levels** of the hydrogen atom, each with its appropriate value of n, the **quantum number,** attached.

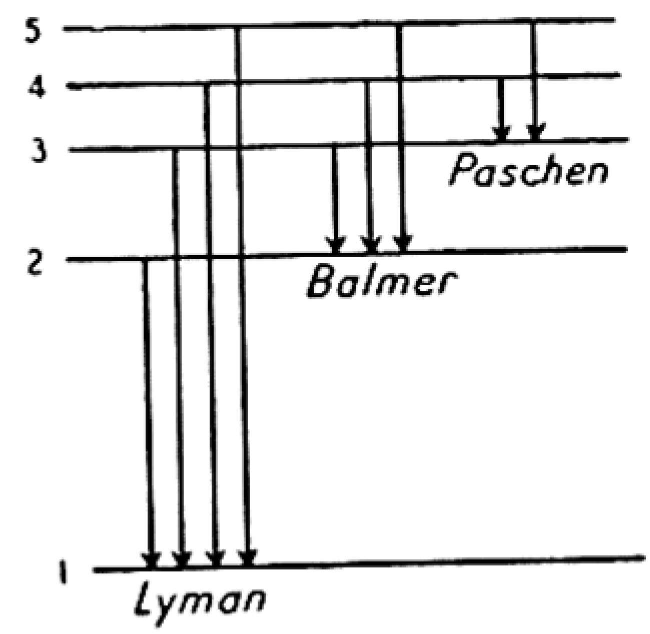


FIG. 746

The remarkable success of the Rutherford-Bohr hydrogen-atom model in explaining the spectral series led to the acceptance of the somewhat revolutionary underlying ideas as to the quantization of the orbits and of the radiation emitted during transitions. The theory was soon modified to take into account various additional effects. One of these is that, since the mass of the nucleus is not infinite, it does not remain stationary while the electron travels round it, and another concerns the fact that elliptical as well as circular orbits are possible. A close examination of spectral lines with apparatus of high-resolving power reveals that each line has a "fine structure," that is to say it is really a group of very nearly equal wave-lengths, and any satisfactory theory of the atom must account for this as well as for the broader details of spectra. We cannot pursue the subject further except to say that the more recently developed wave mechanics has taken the explanation of spectra much further than the original Bohr model could have done. This theory leads to a conception of energy levels within the atom which is more abstract than that of the Bohr orbits.

Absorption Spectra.—When a spectroscope is used to produce a

continuous spectrum of a white-light source, such as the crater of a carbon arc, and a piece of coloured glass is interposed between the source and the spectrum, the glass absorbs some wave-length ranges more than others and the spectrum is modified accordingly. Thus if the glass is red, it is so because it absorbs blue light passing through it, and there will be a dark band in the blue region of the spectrum. The modified spectrum is known as an absorption spectrum and it is characteristic of the absorbing substance. For example, the absorption bands produced by oxygen, nitrogen, etc. can be used to identify these gases. Substances dissolved in water sometimes give absorption spectra by which they can be recognized.

The solar spectrum itself is actually an absorption spectrum. It contains hundreds of dark lines on the background of the continuous spectrum. These are called **Fraunhofer lines,** after the man who made a systematic study of them. Their presence in the solar spectrum can be explained as follows.

The bright sphere of the sun which is normally visible to us emits white light. It is surrounded by an "atmosphere" of cooler vapours known as the **chromosphere,** and it is in this region that the absorption takes place, although a smaller amount occurs also in the earth's atmosphere. The wave-lengths of the various Fraunhofer lines correspond to those in the *emission* spectra of various elements. This correspondence is an example of Kirchhoff's law (page 537, Vol. 2), which states that at a given temperature the emissive power of a body for radiation of a given wave-length is proportional to its absorptive power for the same radiation. In terms of the Bohr theory, we can state that when radiation of all wave-lengths from the centre of the sun falls on an atom in the chromosphere, the orbital electrons are sent out to orbits of higher energy, and the energy absorbed during this process is taken from that part of the radiation which has the same wave-length as is emitted when the electrons fall back to their original orbits. It might at first sight appear that, since the excited atoms subsequently emit the same amount of radiation as they have absorbed, this emission would compensate for the absorption, and the intensity of the light would be undiminished by passage through the absorbing atoms. However, the light reaching the spectroscope from any point in the sun is travelling in a single direction through the chromosphere. Each atom in its path absorbs some of the radiation and then emits an equal amount, but the emission occurs in all directions, so that the amount of light eventually reaching the spectroscope is less than if the radiation proceeded directly without absorption. Hence the lines in the *emission* spectra of the elements in the chromosphere appear as dark lines in the ordinary solar spectrum.

It is an interesting fact that the presence of some Fraunhofer lines whose wave-lengths did not correspond to any emission lines of the elements known at the time of their discovery, was ascribed to a hitherto

undiscovered element which was given the name **helium** and was later discovered on the earth.

A reproduction of the method of formation of the Fraunhofer lines can be set up in the laboratory. The experiment is known as the **reversal of the sodium lines.** Light from an intense white-light source such as the crater of a carbon arc is used to produce a continuous spectrum in a spectroscope. When sodium vapour in a flame or in a heated tube is interposed between the source and the spectroscope slit, two dark lines appear in the yellow part of the spectrum. When the white-light source is removed, the lines appear as the usual yellow D lines of the sodium emission spectrum. The dark lines in the otherwise continuous spectrum are, of course, due to the absorption of light of these wave-lengths by the sodium vapour. They are not completely black—only less intense than the continuous spectrum—because the vapour is itself emitting, otherwise it would not absorb. It is necessary, however, that the temperature of the vapour shall not be high enough to produce sufficient emission to compensate for the absorption.

3. INFRA-RED AND ULTRA-VIOLET

Radiation which is visible to the eye is confined to a small wave-length range extending from roughly 4×10^{-5} cm. (extreme violet) to about $7{\cdot}2 \times 10^{-5}$ cm. (extreme red). Expressed in terms of "pitch," by analogy with sound, the visible range covers less than one octave. Radiation of the same physical character as light can be produced and detected over an enormous range of wave-lengths. Among the shortest known waves are γ-rays. Very roughly their wave-length range may be stated as between 10^{-11} and 10^{-8} cm. Then come X-rays which overlap with γ-rays and extend between about 10^{-10} to 10^{-6} cm. Continuous with this range the ultra-violet (10^{-6} to 4×10^{-5} cm.), which can be produced, in general, by the same methods as visible radiation, brings us to the violet end of the visible spectrum. Beyond the red end of the spectrum we have infra-red radiation which ranges from 10^{-4} to 10^{-1} cm. (*i.e.* 1 mm.), and beyond this are radio waves which are emitted from a circuit in which an electric current is oscillating. The shortest of these (micro-waves) have wave-lengths of the order of 10^{-2} cm. and the range extends upwards indefinitely. The wave-lengths used in radio communication vary from the order of a metre up to kilometres.

There are no gaps in the whole range from short γ-rays to long radio waves. In fact there is a certain amount of overlapping between adjacent categories, except at the ends of the visible spectrum where it is precluded by definition. Thus radio waves can be produced which are shorter than the longest infra-red waves so far detectable. In such a case the description of the waves as "infra-red" or "radio" merely refers to the method of production and detection and not to any inherent difference in their natures.

Infra-red Radiation.—The design of spectroscopes intended for use with infra-red necessarily differs from that of instruments which are concerned with the visible spectrum. In the first place, the detection, observation and measurement of the infra-red spectrum itself, which cannot, of course, be done visually, is carried out either by photography with films specially prepared to be sensitive to infra-red, or else by means of instruments which respond to the heating effect of the radiation, such as the linear thermopile or bolometer. Again, glass strongly absorbs radiation of wave-lengths greater than about 3×10^{-4} cm., so that if the full infra-red spectrum is to be investigated, the prism and lenses of the spectroscope cannot be made of glass. Quartz can be used for longer wave-lengths than glass, and rock-salt (NaCl) or sylvine (KCl) extend the range even further, namely to about 2×10^{-3} cm. These materials have therefore been used to make the prisms of infra-red spectroscopes, and it is possible also to make the lenses of the collimator and telescope of the same substances so as to avoid the absorption which would occur with glass. It is better, however, to substitute concave mirrors for the converging lenses of the visual spectroscope and so to avoid absorption and dispersion altogether. When mirrors are used, the focusing of the instrument may be carried out visually, using a light source and, as reflection is not subject to dispersion, the adjustments so made are valid for the infra-red.

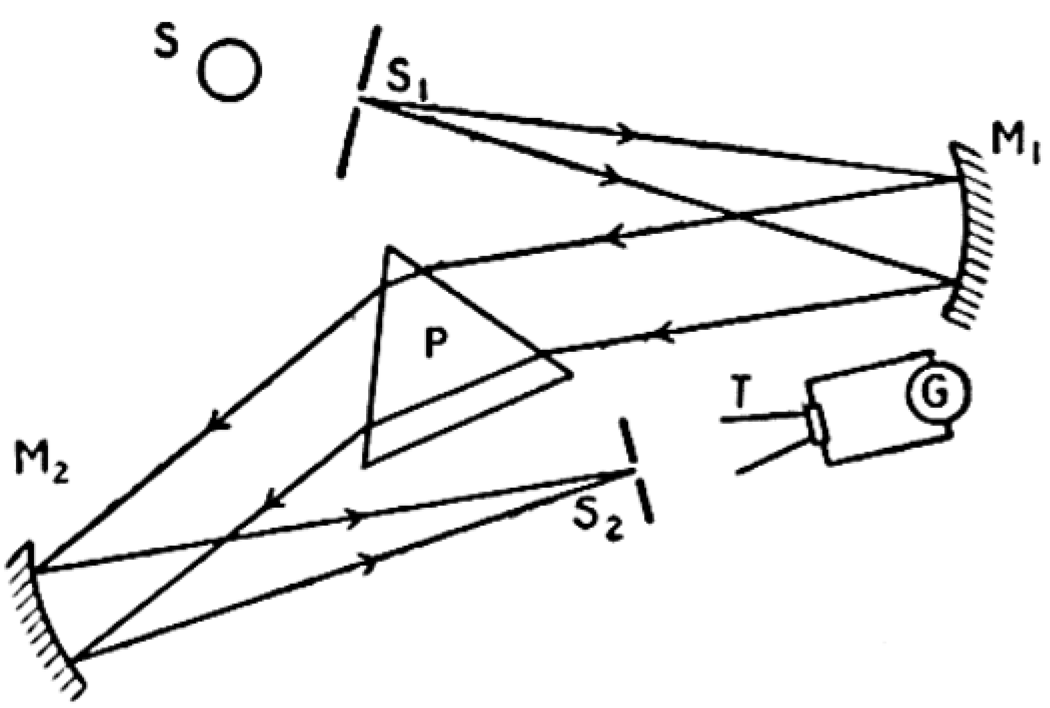


FIG. 747

An infra-red spectrometer using the above principles is shown diagrammatically in Fig. 747. Radiation from the source S passes through the slit S_1, which is situated at the principal focus of the concave mirror M_1 made of polished metal. The parallel rays leaving M_1 strike the prism P, which deviates and disperses them, after which they fall on M_2, which produces a focused spectrum. This may be formed on a screen in which there is a slit S_2 parallel to S_1, in which case S_2 allows the passage of a limited wave-length range of the spectrum, which is then absorbed by the thermopile T and its intensity is measured by the deflection of the galvanometer G. In this way the bands and lines in the infra-red spectrum may be observed. Alternatively a linear thermopile or bolometer may be placed in the plane of the focused spectrum and its response observed as it is moved along the spectrum. Again, as mentioned previously, the spectrum may be recorded on a specially sensitized photographic film or plate.

The wave-lengths of lines and bands in the infra-red spectrum cannot be determined directly by means of a prism spectroscope, but, as with light, this may be done by means of a diffraction grating and by interference methods. A prism spectroscope may be calibrated by means of a source in the spectrum of which the wave-lengths of the infra-red lines have been otherwise determined.

As regards the properties of infra-red radiation, we have already mentioned that, apart from its lack of visual effect, it behaves in a very similar manner to light. For example, it obeys the same laws of reflection and refraction, it is diffracted and can be made to interfere just as light can. When radiation is propagated through a medium, such as the atmosphere, containing many small particles, the direct beam is reduced in intensity by the scattering of the radiation in all directions by the particles. When the particles are smaller than the wave-length, the amount of radiation scattered in any particular direction relative to the direct beam is inversely proportional to the fourth power of the wave-length. Therefore infra-red is scattered less than visible radiation. Thus when photographs of distant scenes are taken through haze due to sufficiently small particles, the direct radiation which the camera lens brings to a focus on the photographic plate is rich in infra-red, so that the clarity of the picture is increased by using plates which are sensitive to infra-red. Before the introduction of radar it was proposed to use infra-red to detect and locate aircraft, etc. at night.

Infra-red radiation is produced by the comparatively low-frequency vibrations of the atoms within molecules, and a study of the wave-lengths present in the infra-red spectrum of a particular substance gives information about the natural modes of vibration of the molecules. Such investigations are made in the infra-red regions of both emission and absorption spectra.

Ultra-violet Radiation.—Glass transmits only that range of ultra-violet wave-lengths which is near the visible violet, and it cannot therefore be used for the prism and lenses of an ultra-violet spectrometer. The commonest material used is quartz. Mirrors are not used to replace the lenses, as they are in infra-red spectroscopy, because the intensity of reflection of the ultra-violet is small. Diffraction gratings are used for actual wave-length measurements. The ultra-violet spectrum can be made visible by focusing it on a screen coated with some substance which fluoresces, *i.e.* emits visible light, when ultra-violet falls on it. By far the commonest method, however, is to focus the spectrum on to a photographic plate or film, the ultra-violet being strongly actinic. In many spectroscopes designed for use in the ultra-violet the instrument is totally enclosed and evacuated of air to eliminate absorption from this cause. These are known as **vacuum spectrographs.**

The proportion of ultra-violet in the thermal radiation from a hot solid body is comparatively small. It is much greater in the radiation

from electric arcs struck between electrodes of metals such as iron and tungsten. The flame of the arc consists of vapourized metal at a very high temperature. A mercury arc is frequently used as a source of ultra-violet. It consists of an arc discharge between two pools of mercury in a closed vessel. The principle of the **mercury vapour lamp,** as it is called, is illustrated in Fig. 748. The vessel, which is evacuated and is made of quartz to withstand the high temperature and to minimize ultra-violet absorption, consists of a tube joining two bulbs. Wires are fused through the wall of each bulb so that the two pools of mercury may be connected to an electrical potential difference. When the lamp is carefully tilted, a thread of mercury connects the two pools and the current flows. The thread is broken by restoring the tube to its horizontal position, when the arc strikes and is maintained. There are several spectral lines in the visible part of the spectrum of such a lamp, as well as in the ultra-violet, and it is sometimes used as a source of monochromatic light, the unwanted visible lines and the ultra-violet being removed by passing the radiation through a suitable filter.

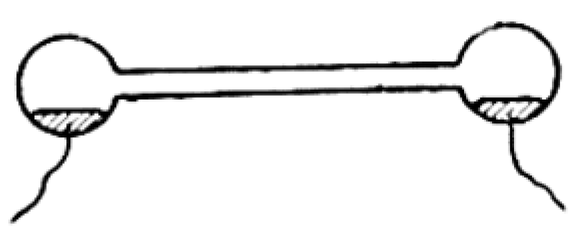


FIG. 748

The effects of ultra-violet radiation on living things such as plants, animals and human beings are well known. Its tonic effect on the human body is due to the fact that, although the ultra-violet can penetrate the skin to only a very small depth, it nevertheless takes part in the production in the body of vitamin D, which is essential to healthy life and, among other things, to the avoidance of the bone disease known as rickets in young children. The radiation responsible for this process has a wave-length in the neighbourhood of 3000 Å and is present in sunlight. Much of the ultra-violet present in the actual radiation from the sun is absorbed in the atmosphere (particularly if smoke and dirt are present), so that the amount which actually reaches us increases with the elevation of the sun. If "artificial sunlight" is used as a substitute it is necessary to ensure that the radiation from the lamp (*e.g.* a mercury vapour lamp) is within the correct wave-length range. Too large a dose of natural or artificial sunlight is, of course, exceedingly harmful.

Ultra-violet, especially of short wave-length, kills bacteria in the air or in substances which it can penetrate, and it is sometimes used to maintain a sterile atmosphere where this is necessary in laboratories or factories.

Ordinary photographic films and plates are very sensitive to ultra-violet radiation and, because of its short wave-length, it is strongly scattered by the atmosphere. If a photograph is being taken of a distant mountain or clouds, the atmosphere between the scene and the camera scatters ultra-violet and short wave-length light from the sun into the lens. The result is a suffusion of short wave-length radiation over the whole distant scene, which produces over-exposure of the negative and lack of sharpness. The phenomenon is similar to the general

illumination which is seen surrounding a street lamp viewed in fog, and it can be avoided by placing a filter in front of the camera lens to reduce the intensity of the ultra-violet and of the visible violet and blue which also contribute to the effect.

When a substance becomes self-luminous as a result of radiation falling upon it, it is said to be **fluorescent** or **phosphorescent** according as its luminosity ceases at the instant when the incident radiation is cut off or whether it decays more gradually. In general, fluorescence is exhibited by liquids and vapours and phosphorescence by solids, the luminous paint on watch dials being a common example of the latter. A rule known as "Stokes' law" states that the fluorescence or phosphorescence does not contain wave-lengths which are shorter than those in the exciting radiation. In accordance with this the phenomena are frequently observed when substances are irradiated with ultra-violet. Many minerals emit characteristic colours when placed in a beam of ultra-violet radiation and, in particular, precious stones can be identified and imitations detected by this means.

4. COLOUR

At the outset of this short discussion of what is in fact a very complicated subject, it is essential to realize that the assessment of colour is a subjective matter. We know that the colour sensation produced by monochromatic light depends upon its wave-length, and people with normal vision can usually agree about the names of the colours associated with the different wave-lengths in the spectrum. Colour blindness is a common defect, however, and those who suffer from it receive quite a different impression of colour compared with that experienced by people with normal colour vision. Sometimes colour blindness is so extreme that a sense of colour is entirely lacking. It is thought that animals have little, if any, ability to discriminate between colours.

In a discussion of the subjective effects of sound (Chapter XXXVIII, Vol. 3) it is stated that, except for the small class of people who are tone deaf, we can discriminate between the musical pitches of pure notes of different frequencies. This is analogous to our ability to appreciate and judge the colours of monochromatic light of different frequencies, and we ought to mention once again that it is the frequency of monochromatic light which fixes its colour and not primarily the wave-length, which depends upon the medium in which it is travelling.

Again, when we listen to two or more pure notes simultaneously, or when we hear a single note and its overtones, we experience a sensation distinct from pitch which we call the "quality" of the sound, and which depends upon the frequencies present in the vibrations stimulating the ear and on their relative intensities. At the same time our auditory equipment has the ability to recognize as separate ingredients the individual notes or overtones which are present in any particular mixture.

This latter faculty has no visual counterpart. When two or more monochromatic lights are used to illuminate simultaneously the same area of a screen, we see a single patch of colour which is, of course, determined by the separate colours, but which is different from each of them, and the existence of the separate colours is quite undetectable. The eye and brain, so to speak, mix the colours and register the result as a distinct and different colour.

The Mixing of Coloured Lights.—In discussing colour mixing it is essential to distinguish between the mixing of lights, which means the stimulation of the retina simultaneously by lights of different wave-lengths, and the mixing of pigments as in painting. We shall deal with the mixing of lights first.

Suppose that we set up an arrangement (Fig. 749) for producing a focused spectrum VR of a white-light source using two lenses as in

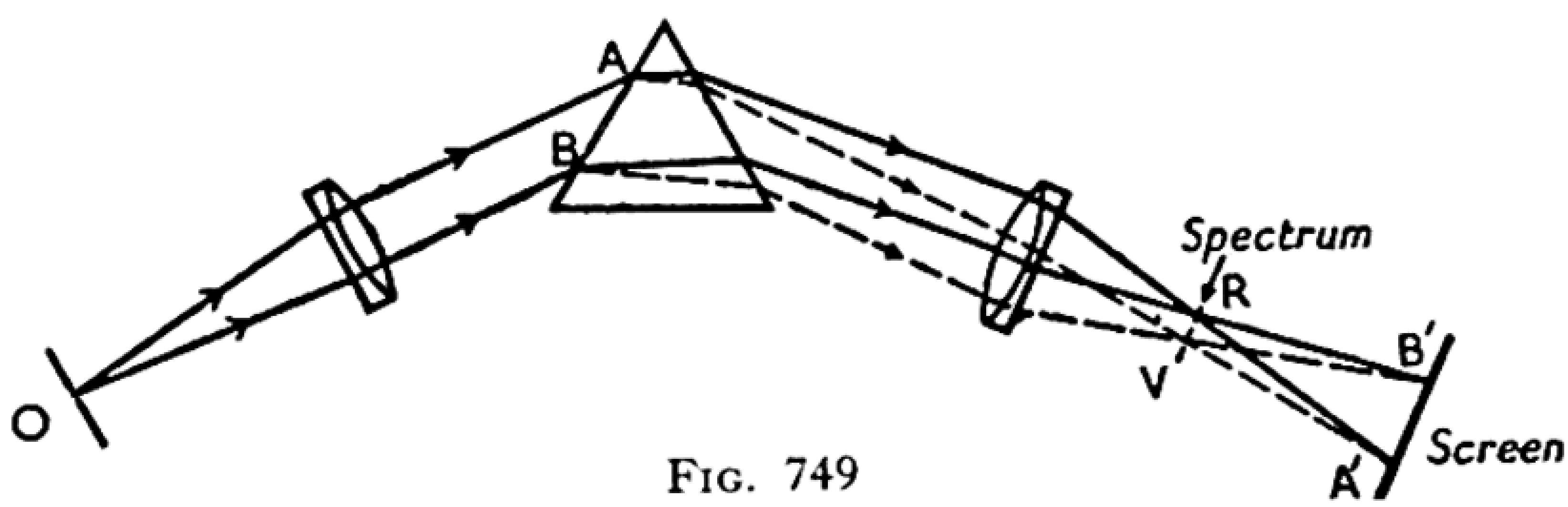


FIG. 749

Fig. 740, and that, instead of receiving the spectrum on a screen, we let the rays pass through it. At a certain distance beyond the spectrum the coloured rays diverging from a point such as A or B on the first face of the prism are brought to a common focus, *i.e.* they are recombined to form white light. Therefore a screen placed in this position receives a patch of white light which is, in fact, an image of the first face of the prism. This can be verified by placing a pin against the prism face, when its image will be formed on the screen.

Many experiments on colour mixing have been performed with apparatus of the sort just described, which is sometimes called a "colour-patch" apparatus. For example, a thin straight obstacle, such as a piece of wire, may be placed in the plane of the spectrum and at right angles to its length so as to intercept, according to its position in the spectrum, the light within a small wave-length range depending upon the width of the obstacle. The light patch on the screen is then coloured, being white light deprived of the wave-length range cut off by the obstacle. In other words, it is a mixture of the remaining unobstructed spectral colours. The colour of the patch depends, of course, on the position of the obstacle. A pair of coloured lights are said to be **complementary** if together they form white light, so that in this experiment the patch of light on the screen is the complementary colour to the portion of the spectrum which is blocked out.

Again, we can place in the plane of the spectrum a screen fitted with two slits (each perpendicular to the length of the spectrum) whose positions and widths are separately adjustable. With this arrangement the two portions of the spectrum which pass through the slits are superposed on the screen S, and so we are able to determine which pairs of spectral colours make the screen appear white, *i.e.* are complementary. If the slits are narrow, the light passing through each is approximately monochromatic, and their intensities can be varied by adjusting the widths of the slits, although this also has some effect on the wave-length range transmitted.

When one slit, say S_1, is placed near the extreme red end of the spectrum, it is possible to make the illumination of the screen S appear white by placing the other slit S_2 in the green-blue part of the spectrum and adjusting their relative widths. Thus the complementary colour to long wave-length red is green-blue. If S_1 is then moved through the red towards the middle of the spectrum, S_2 must be moved towards the violet end, *i.e.* in the same direction as S_1, in order to preserve the white illumination on the screen. Thus the whole series of complementary spectral colours can be recorded. Proceeding in this way, we find that S_2 has moved off the violet end of the spectrum before S_1 has reached the green portion, which means that green has no spectrally pure complementary colour. Its complementary colour is, in fact, a mixture of blue and red, which is known as magenta.

When we mix two coloured lights which are too close to each other in the spectrum to be complementary, the resulting colour lies somewhere between them in the spectrum.

A further important series of observations can be made with an arrangement such as that in Fig. 749 if we place *three* slits of adjustable width in fixed positions in the spectrum. When the slits are narrow, the light passing through each is approximately monochromatic, and the patch of light on the screen S is then a mixture of three separate wave-lengths which are determined by the positions of the slits. Their relative intensities are determined by the relative widths of the slits. When the slits are widely separated in the spectrum so as to transmit wave-lengths in the red, green and blue respectively, it is found that merely by altering their relative intensities it is possible to produce a very large range of colours on the screen, including white light. The three separate colours are called **primaries.** It is not true to say that *all* colours can be imitated by mixing various amounts of three primaries unless we allow negative amounts of one of the primaries in certain cases. By a negative amount of one colour we mean that the appropriate intensity of this primary is added to the colour which is to be imitated and we then match the result with a mixture of the two remaining primaries. The three widely separated primaries mentioned above require the use of negative values to a smaller extent than if the chosen primaries are nearer each other in the spectrum.

In the science of colour measurement use is made of the fact that every colour can be produced by a mixture of primaries in order to specify quantitatively any given colour in terms of the relative intensities of a given set of primaries which are required to match it.

Colour by Absorption.—A common example of colour production by absorption is the action of stained glass. When white light passes through coloured glass, certain components in the spectrum are absorbed more than others, and the "colour" of the glass is due to the mixing, in the eye, of what is left of the original white light. It is not necessarily true to say that red glass "absorbs all colours except red." Naturally, red is predominant in the light emerging from the glass, but it is not spectrally pure red. Coloured transparent filters which are made for use, for example, in lens experiments where chromatic aberration is to be avoided, do not give truly monochromatic light when white light is passed through them. They can, however, be made to suppress all spectral lines except a chosen one when they are placed in front of a source, such as a mercury vapour lamp, which gives a spectrum of fairly widely separated lines.

When white light is passed in turn through two pieces of differently coloured glass, the emergent light consists of those wave-length ranges which are not completely absorbed by either, and if there is no overlapping of the wave-lengths transmitted by each, then no light at all will get through the pair.

The production of colour by absorption is described as a **subtractive** process in contrast to the **additive** process of mixing different coloured lights together which we have discussed in the first part of this section.

When an object is illuminated by white light and viewed through coloured glass, its colour appears unnatural because of the removal by absorption in the glass of some of the colours which the object normally sends out.

It should be realized that the absorption in transparent materials such as coloured glass occurs in the material itself after the light has entered it. Light which is reflected from the surface without entering the material is unaffected by the absorption. Thus, in Fig. 750, the white light reflected from the top surface of a block of red glass remains white. The light which enters the glass and follows the dotted path emerges as red light because of the selective absorption which it experiences in the glass.

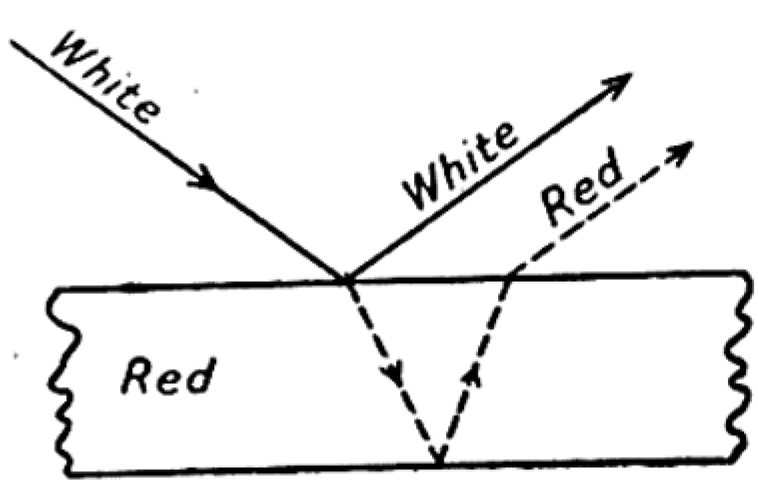


FIG. 750

Pigments.—Although it might not appear so at first sight, the colours of pigments such as are used in painting and printing are due to the same sort of effect as is illustrated in Fig. 750. The coloured pigment particles are transparent, and are suspended in a transparent medium called the "vehicle." Light which falls on a layer of pigment is not very strongly

reflected at the surfaces of the particles, because the latter are surrounded by a medium whose refractive index is not very different from their own, and the strength of a reflection increases with the difference between the refractive indices on the two sides of the reflecting surface. Much of the light striking a layer of paint, therefore, penetrates below the surface of the layer before it is reflected, and it suffers selective absorption as it passes through the pigment particles, finally emerging with its characteristic colour. The reflected light leaves the layer in all directions because of the haphazard arrangement of the surfaces of the pigment particles at which the reflections take place. The reflection is therefore diffuse and not specular (*i.e.* mirror-like). Obviously a layer of pigment loses its natural colour when it is illuminated by other than white light. For example, red paint appears almost black in blue light because its normally red colour is due to its strong absorption of blue light falling on it.

The production of new colours by mixing pigments is a subtractive process. The usual example given is the mixing of blue and yellow paints to form green. This is possible because green is the colour which is not absorbed by either the blue or the yellow pigment particles.

Colour Vision.—We shall not attempt to give any explanation of how, when monochromatic light enters our eyes, we see it as a definite colour, and of how two or more monochromatic lights falling simultaneously on the same region of the retina give the impression of a colour which is different from the separate colours. A complete and satisfactory theory of colour vision has not yet been established, but we may note that the fact that nearly all colours can be produced by the superposition of three monochromatic primary colours in varying proportions has led to the **three-colour theory.** This supposes that the eye possesses three separate receptor mechanisms which, when stimulated, give the sensations of red, green and blue respectively. It is supposed that each is stimulated by light of any wave-length, but that the maximum response of the red mechanism occurs when the incident light is red, and similarly for the other two colours. Thus, when light of any wave-length enters the eye, each mechanism responds to a certain definite extent, and the brain is supposed to interpret these three simultaneous responses as a single colour. On this theory, colour blindness is due to a defective response to one or more of the primary colours. There are difficulties in the way of complete acceptance of the three-colour theory of colour vision, however, but we shall not discuss any other theories or give any more facts, although a large amount of information has been discovered about normal and defective colour vision.

5. THE RAINBOW

The rainbow is a naturally occurring impure spectrum seen when the sun shines on falling rain. In order to discuss its formation it is first

necessary to consider the action of a spherical water-drop on a ray of light as shown in Fig. 751. The ray enters the drop at A, where it is refracted, after which it is partially reflected at B and refracted at C. If the angles of incidence and refraction at A are i_1 and i_2 respectively, it is easily seen that the other angles have the values indicated in the drawing. The angle through which the original incident ray is deviated by the drop is δ, which can readily be shown to be related to i_1 and i_2 by the equation

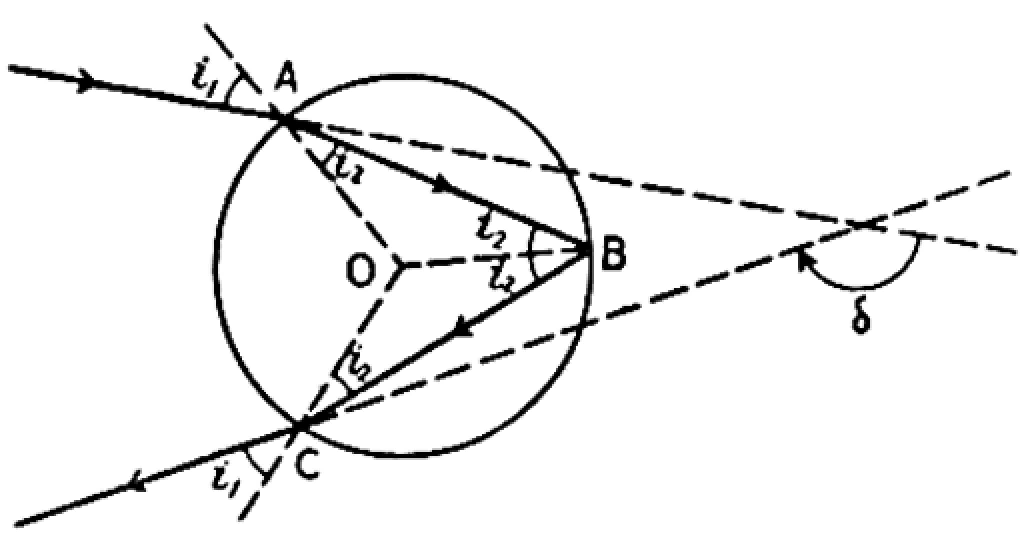


FIG. 751

$$\delta = 180° + 2i_1 - 4i_2$$

Evidently it is possible to use this equation together with Snell's law (to find i_2) in order to construct a graph of δ against i_1 for a given value of the refractive index of the drop. When this is done it is found that δ has a minimum value, which in the case of a water-drop is about 138° for red light and 140° for violet light.

Fig. 752 shows the action of a drop of water on a series of parallel incident rays. Ray (3) is intended to represent the ray for which the deviation is a minimum, and rays (1) and (5) are the two extremes on either side of minimum deviation. Thus ray (1), which enters and leaves the drop tangentially ($i_1 = 90°$), undergoes the maximum possible deviation for rays entering the drop above ray (3), while ray (5) ($i_1 = 0°$), which returns along its own path, suffers the maximum deviation (180°) of all the rays below ray (3). In order to avoid complication of the ray diagram inside the drop, only the incident and emergent directions of the intermediate rays (2) and (4) are shown. For parallel rays incident on the lower half of the drop there will be corresponding rays emerging from the top half symmetrically with those drawn in Fig. 752.

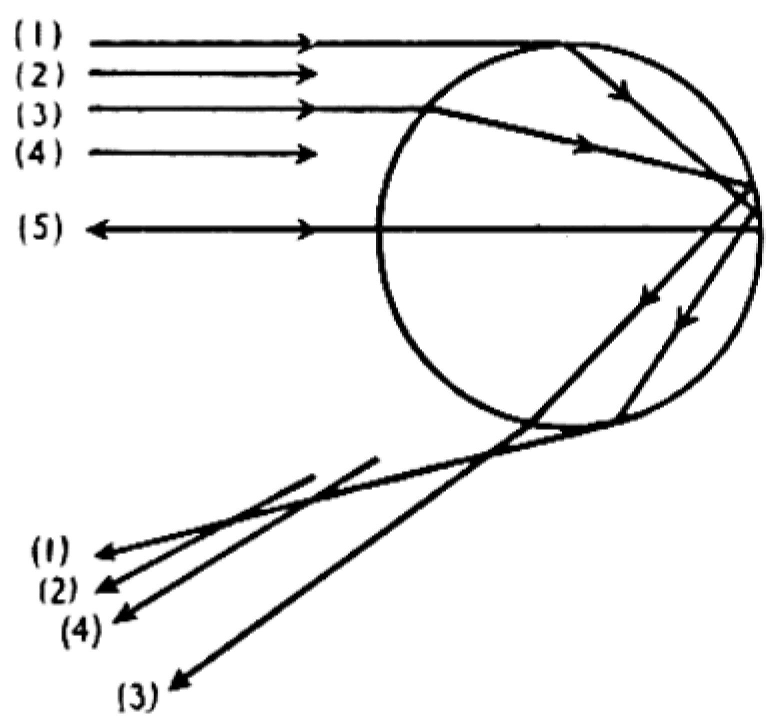


FIG. 752

In the neighbourhood of the minimum deviation the direction of the emergent ray changes only slowly from one ray to the next. That is to say, the emergent rays such as (2), (3) and (4) are approximately parallel. This means that an eye looking at the drop along the direction of ray (3) receives a greater proportion of the light contained in the incident rays (1) to (5) than if the drop is viewed from any other direction. This concentration of light in one direction is the cause of rainbows. The colouring

is due to the dispersion which occurs at each refraction and causes the direction of maximum light to vary with wave-length.

We can now discuss the formation of a rainbow under circumstances depicted in Fig. 753. The direction of the parallel rays from the sun is shown by two arrows. An observer E viewing the raindrops with his back to the sun will receive the greatest amount of light from those drops whose positions are such as to send to E the rays which they deviate by the minimum amount. For red light this minimum deviation is about 138°, so that the drops which send the most red light to E must lie on straight lines such as EA and EB which are inclined at an angle θ to EX

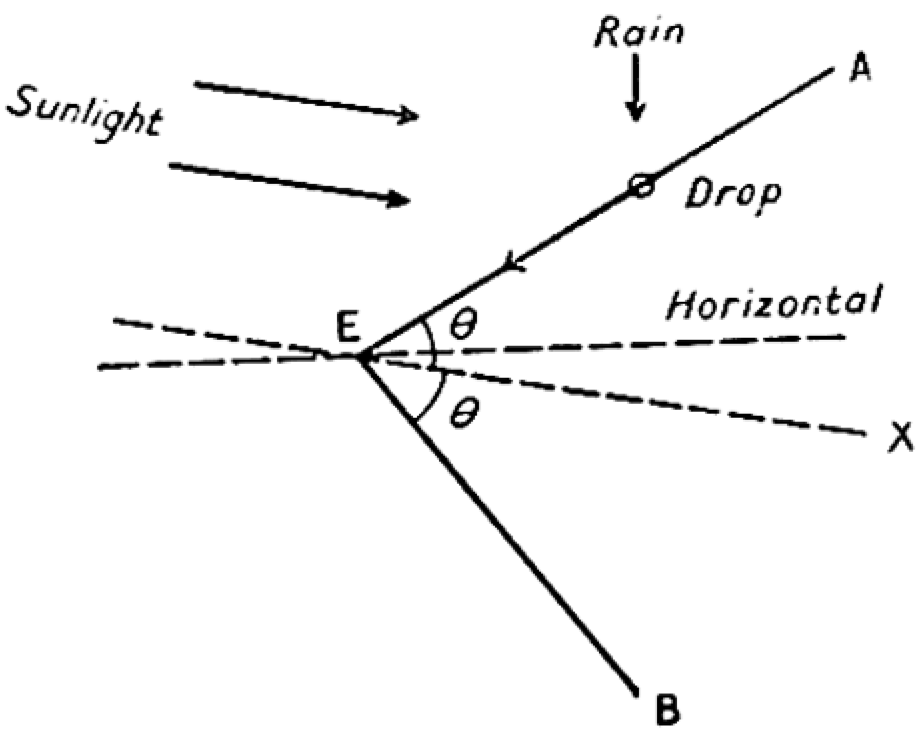


FIG. 753

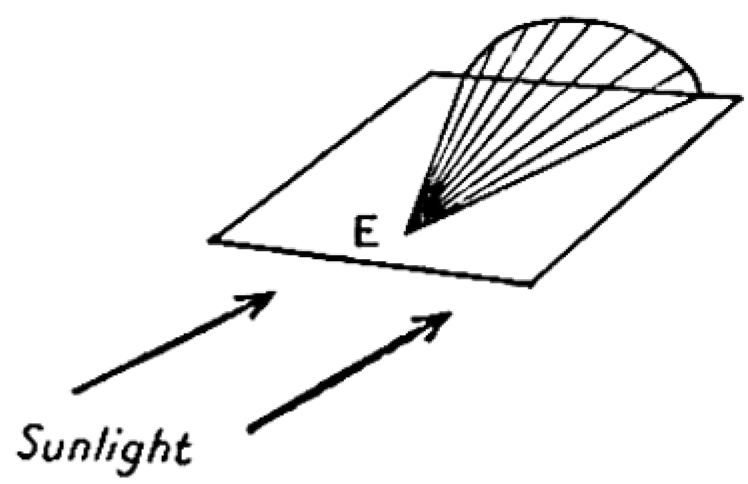


FIG. 754

(which is parallel to the sun's rays), where θ is equal to (180°–138°), *i.e.* 42°. There are an infinite number of such lines, and together they form the surface of a cone of semi-angle θ with its apex at E and its axis along EX. All drops which send the maximum amount of red light to E therefore lie on the surface of this cone. Drops which send the maximum amount of violet light lie on another cone having the same apex and axis but a slightly smaller value of θ, namely (180°–140°), *i.e.* 40°. The observer therefore sees a spectrum in the shape of a circular arc with the red on the outside edge. A certain amount of light reaches the observer from drops inside the arc, the deviation for the drops in such a position being greater than the minimum, but no light can be received from drops outside since this would involve a deviation less than the minimum. We shall mention later, however, that a double reflection in each drop can give rise to a bow outside the primary bow.

From Fig. 753 it can be seen that the complete circle of the rainbow is visible at E only if the observer is in a sufficiently elevated position to allow the existence of the complete conical surface within the shower of rain. On level ground, however, as depicted in Fig. 754, only part of the circle is visible. It is easy to see that if the sun's rays were horizontal half of the cone would be above ground-level and the bow would be a semicircle. As the elevation of the sun increases, less of

the bow remains visible, until at an elevation of about 40° it cannot be seen at all.

A rainbow formed in the way just described by a single reflection inside each drop is called a **primary bow.** It is possible for light to be sent back to an observer from a shower of rain by a process involving two reflections in each drop as shown in Fig. 755. The deviation δ is greater than 180° in this case, and the average minimum deviation for all colours is about 233°, thus making the acute angle between the incident and emergent rays equal to about (233°–180°), *i.e.* 53°. Therefore if the light is strong enough to be visible after the two partial reflections, it will reach the observer from drops on the surface of a second cone of semi-angle about 53°. This constitutes the **secondary bow** which is often seen. On account of the crossing over of the incident and emergent rays the acute angle between them is greater for violet than for red light. This causes the violet to be on the outside of the secondary bow, whereas it is on the inside of the primary bow. Light can reach the eye from points outside the secondary bow because this involves deviations greater than the minimum, but it is not possible from points inside the secondary bow. The region between the primary and secondary bows cannot send light to the eye by either one or two reflections, and it consequently appears darker than the rest of the sky in the neighbourhood of the bows.

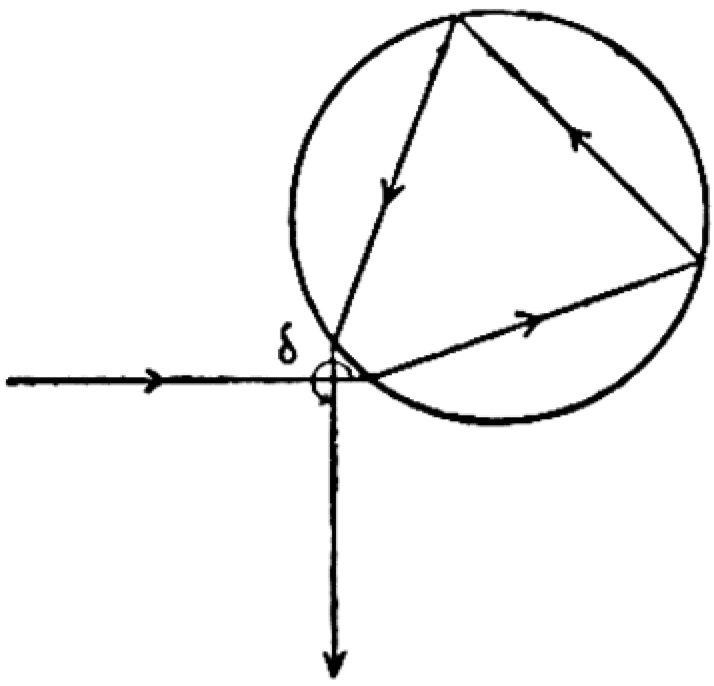


FIG. 755

Diffraction effects are often superimposed on the purely dispersion effects described above, and give rise to "supernumerary" bows near the edges of the main rainbows.

EXAMPLES LIII

1. Draw a labelled diagram of a spectrometer set up for studying the deviation of light through a triangular prism. Describe how you would adjust the instrument and use it to find the refractive index of the prism material.

Indicate briefly how you would show that the radiation from an arc lamp is not confined to the visible spectrum. (L.H.S.)

2. Describe and give a diagram of the optical system of a spectrometer. What procedure would you adopt when using the instrument to measure the refractive index of the glass of a prism for sodium light? What additional observations would be necessary in order to determine the dispersive power of the glass?

The refractive index of the glass of a prism for red light is 1·514 and for blue light 1·523. Calculate the difference in the velocities of the red and blue light in the prism if the velocity of light *in vacuo* is 3×10^5 kilometres per sec. (J.M.B.H.S.)

3. Draw a graph showing, in a general way, how the deviation of a ray of light when passed through a triangular prism depends on the angle of incidence.

You are required to measure the refractive index of glass in the form of a prism by means of a spectrometer provided with a vertical slit. Explain how you would level the spectrometer table and derive the formula from which you would calculate the refractive index. (You are not required to explain any other adjustment of the apparatus nor to explain how you find the refracting angle of the prism.) (L.I.)

4. Explain what is meant by a line spectrum. Describe any one form of source, such as a hydrogen discharge tube or mercury vapour lamp, which gives a line spectrum.

Describe carefully how you could obtain a beam of monochromatic light from such a source, and how you would measure the wave-length of this light. (O.H.S.)

5. Give an account of the solar spectrum. Point out and explain as fully as you can the chief differences between this spectrum and that of an ordinary tungsten filament lamp.

Explain, treating one chosen example fully, why coloured fabrics which match in artificial light are often found not to match in daylight. (O.H.S.)

6. (*a*) What is meant by *chromatic aberration*? Describe an experiment to demonstrate it, and explain how it may be reduced by cementing two lenses together.

(*b*) Explain briefly the causes of the colours of (i) green leaves, (ii) the rainbow, (iii) soap bubbles. (J.M.B.H.S.)

7. Write a short essay on colour, and explain clearly why two objects which appear exactly the same colour by gaslight may appear of different tints by daylight. (L.Med.)

Chapter LIV

POLARIZED LIGHT

1. PRODUCTION AND NATURE OF POLARIZED LIGHT

Polarization by Reflection.—In Fig. 756 (i), M_1 and M_2 are two plane reflectors made of, say, black glass, the purpose of the blackening being to confine reflection to the front surface. A ray of light incident on M_1 is reflected to M_2, where it is reflected again. A remarkable observation can be made with this arrangement, namely that if M_2 is rotated about the direction of the ray incident upon it, *i.e.* in such a way that the angle

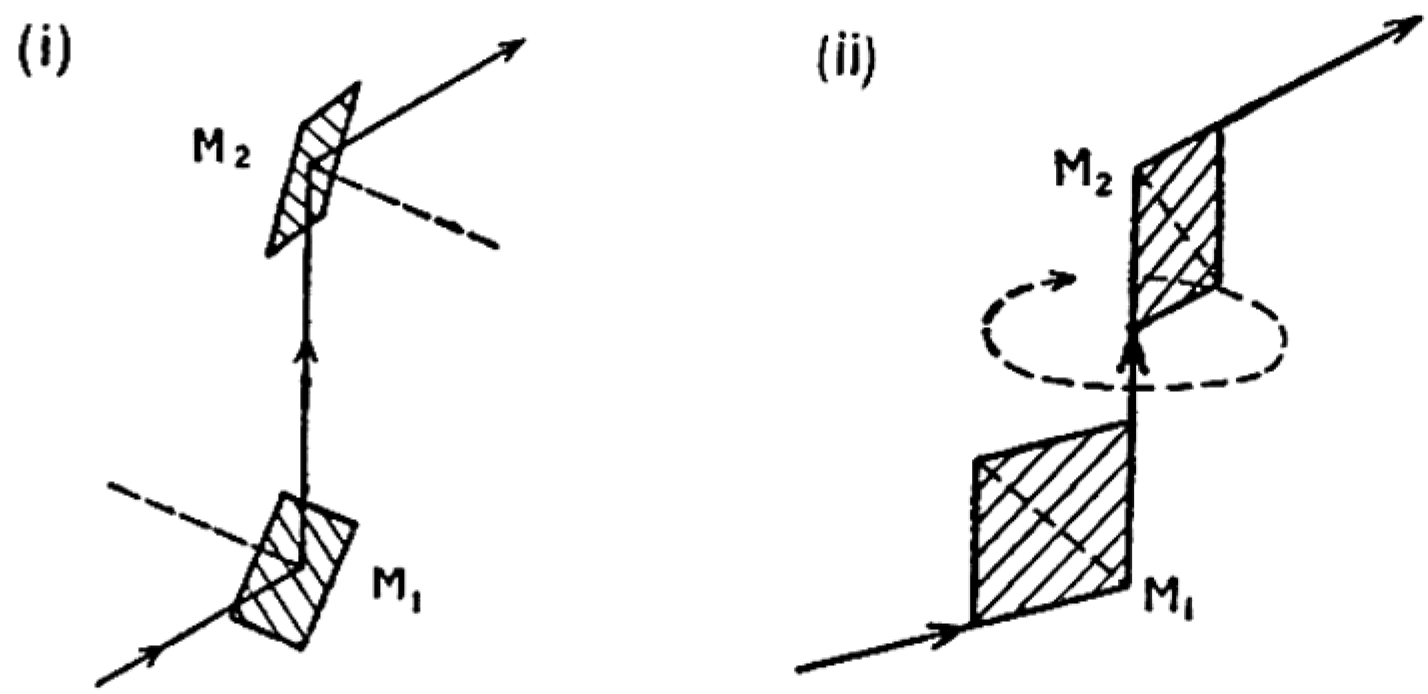


FIG. 756

of incidence upon it is constant, then the intensity of the light reflected from it changes continuously. The intensity passes through two maxima and two minima during each complete revolution. The intensity of the minima depends upon the angle of incidence of the light on both M_1 and M_2. In the case of glass, the minimum intensity is practically zero when the angle is about 57° in both cases. We then have the rather surprising fact that it is possible for two successive reflections to result in practically no light at all.

The significance of the above observations lies in the fact that the rotation of M_2 about the direction of the light incident upon it involves the rotation of the plane of incidence on M_2 relative to that on M_1. This is illustrated in Fig. 756 (ii), from which the actual mirrors have been omitted, and in which portions of the two planes of incidence are shown as shaded areas with the normals as diagonals. As the top mirror rotates about the ray M_1M_2, the plane of incidence on it rotates as indicated by

the dotted arrow. The light reflected from M_2 has its maximum intensity when the two planes of incidence are parallel to each other, and its minimum intensity when they are mutually perpendicular. Thus there are two minima and two maxima for each complete revolution of one plane of incidence with respect to the other.

One obvious way of explaining this effect is to say that the first reflection gives some property to the reflected light in a certain direction at right angles to its direction of propagation M_1M_2. Such an asymmetry is called **polarization,** and in the absence of other evidence as to what the certain direction is, we can assume it to lie in the plane of incidence on M_1. We then describe the reflected light as "plane polarized in the plane of incidence." In other words, we say that the plane of polarization of the reflected light is the same as the plane of incidence. This was the original method of visualizing polarization.

When we bring the wave nature of light into consideration we can associate polarization with the direction of vibration of the "particles" taking part in the wave motion. Thus we picture a beam of ordinary light as a collection of *transverse* waves in which no particular direction of vibration is favoured more than another. When the beam is plane polarized, however, the vibrations take place in only one of the infinite number of planes containing the direction of propagation. It is obvious therefore that, whatever method is used for polarizing a beam of light (and we shall discuss others later), the intensity of the light is diminished in the process, because for any given vibration present in the original beam only the component in the permitted direction is retained in the polarized beam.

It will immediately be realized that this explanation of polarization in terms of the direction of vibration precludes the possibility of light being a longitudinal wave motion, because if the vibrations took place along the direction of propagation there would be no property of the wave in the transverse direction which could be modified. The existence of polarization, therefore, provides evidence that light consists of transverse waves. It so happens that experimental observation indicates that the plane in which the vibrations are occurring in a beam of polarized light is perpendicular to the plane of polarization as originally defined, *i.e.* to the plane of incidence in the case of polarization by reflection. Some people have therefore adopted the practice of referring to the plane in which the vibrations take place as the plane of polarization. We shall keep to the original nomenclature however, and in order to avoid having to keep in mind the fact that the vibrations are *perpendicular* to the plane of polarization as we define it, we shall speak more frequently in terms of the **plane of vibration** than the plane of polarization. It is almost unnecessary to state that the unaided eye cannot normally detect the polarization of light.

Our explanation of the polarization of light by a mirror is therefore as follows. The light incident on the first mirror M_1 consists of transverse

waves in which the vibrations are taking place in all directions perpendicular to the direction of propagation. Let each vibration be resolved into a component in the plane of incidence and another in the perpendicular direction. The effect of the reflection is in general to reduce the components in the plane of incidence (the plane of polarization), thus giving a greater proportion of vibrations in a perpendicular direction (the plane of vibration).

It is found that the extent of the polarization (*i.e.* the degree to which vibrations in the plane of incidence are eliminated) varies with the angle of incidence and is maximum when, as in Fig. 757, the angle of incidence is such as to cause the refracted and reflected rays to be perpendicular to each other. The vibrations of the light waves in the plane of the drawing are indicated in the diagram by the short transverse lines and those in a perpendicular direction are indicated by dots. Vibrations in all transverse directions are present equally in the incident light, but only the two just mentioned are shown in the diagram. The angle of incidence for maximum polarization is called the **polarizing angle i_p**, and it is easy to see from Fig. 757 that the angle of refraction in the reflecting medium is equal to $(180° - 90° - i_p)$, *i.e.* $(90° - i_p)$. Snell's law applied to the refraction gives, therefore,

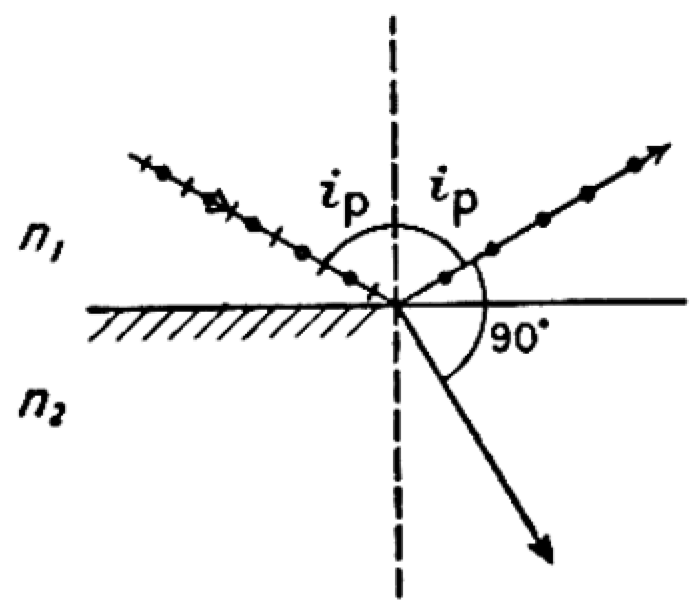


FIG. 757

$$n_1 \sin i_p = n_2 \sin (90° - i_p)$$

$$= n_2 \cos i_p$$

$$\therefore \quad \frac{n_2}{n_1} = \tan i_p$$

Thus the polarizing angle is the angle whose tangent is equal to the refractive index of the reflector relative to the medium in which the incident light is travelling. This is known as **Brewster's law.** For light incident on a glass surface in air, we have $n_1 = 1$ and $n_2 = 3/2$ approximately, so that $\tan i_p = 3/2$, which gives $i_p = 57°$ approximately.

In the experiment with the two mirrors, the first mirror, which polarizes the light, is called the **polarizer,** and the second mirror, by means of which the polarization is made evident, is called the **analyzer.** When light is incident on the polarizer at the polarizing angle, the light which leaves it is almost completely plane polarized, *i.e.* the vibrations are almost entirely confined to a direction perpendicular to the plane of incidence. The second mirror, the analyzer, would, of course, have the same effect on a beam of ordinary light incident at the polarizing angle. Let the polarized beam from the polarizer be incident at the polarizing angle on the analyzer.

In Fig. 758, the vector I represents the amplitude and direction of the transverse vibration of the light coming from the polarizer, the direction of propagation of the light being perpendicular to the paper. The line OV making an angle θ with I represents the direction of the vibrations which the analyzer will reflect, *i.e.* it is perpendicular to the plane of incidence on the analyzer. This latter plane is represented by OP, and vibrations in this direction are stopped by the analyzer. The component of I in the direction OV is $I\cos\theta$ and represents the amplitude of the waves reflected by the analyzer. The intensity of the light leaving the analyzer is proportional to the square of the amplitude of the vibration, *i.e.* to $I^2\cos^2\theta$. The proportionality of the intensity to $\cos^2\theta$ is known as the **law of Malus** and is found to be verified by experiment.

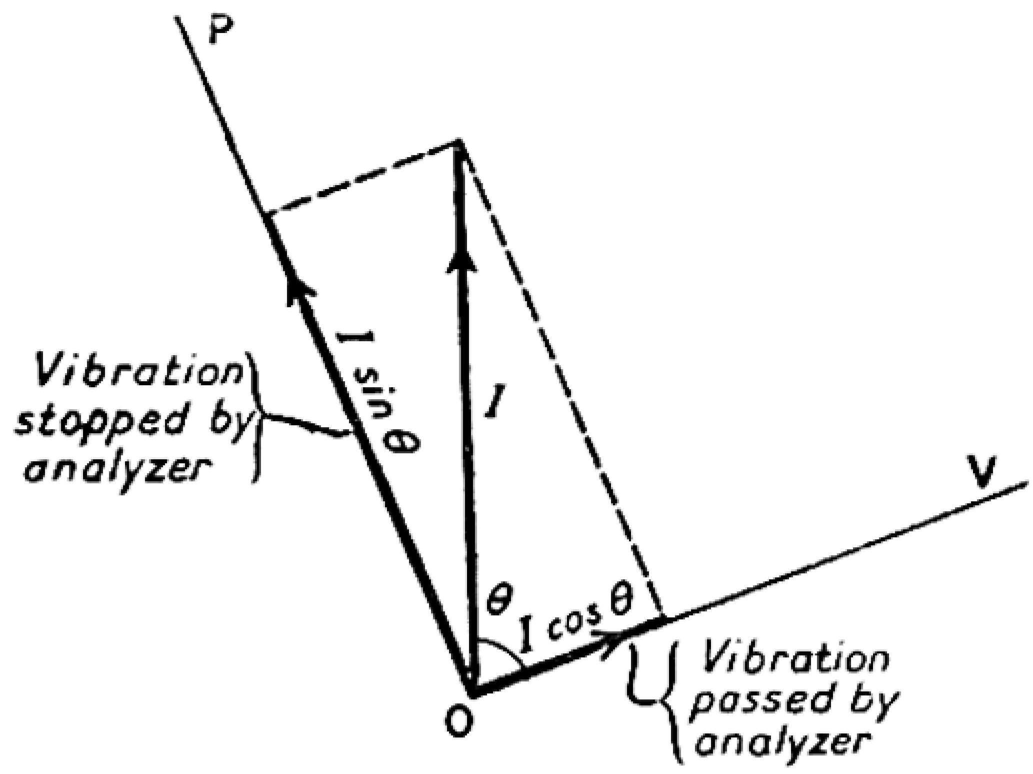


FIG. 758

The light leaving the analyzer has a maximum intensity when $\cos^2\theta = 1$, *i.e.* when $\cos\theta = \pm 1$ and $\theta = 0$ and $180°$. This occurs when the planes of incidence on the polarizer and analyzer are parallel to each other. Midway between these two positions, *i.e.* when $\theta = 90°$ and $270°$, the intensity is zero.

We now consider briefly the state of polarization of the light which is refracted into the material of a plane reflector. It must be remembered

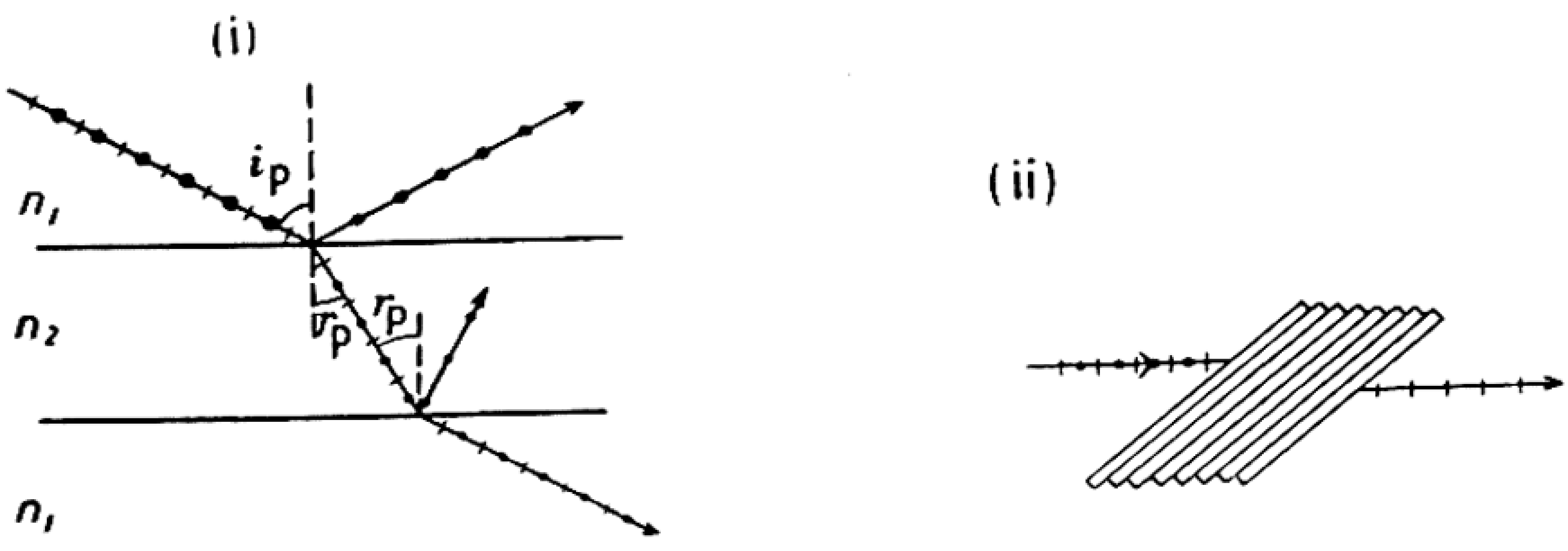


FIG. 759

that reflection is only partial, so that the fact that at the polarizing angle the vibrations in the reflected beam are completely confined to a direction perpendicular to the plane of incidence does not mean that there are no such vibrations present in the transmitted beam. The fact is, rather, that the transmitted beam is poorer than ordinary light in vibrations perpendicular to the plane of incidence since these are the most strongly reflected. It is therefore partially polarized. When, as in Fig. 759, light is incident on

one surface of a parallel-sided slab at the polarizing angle, it is easy to show that the refracted light strikes the far side of the slab at an angle (r_p) whose tangent is n_1/n_2 and which is therefore the polarizing angle for internal reflection. This means that the light which is internally reflected is plane polarized, and that the light which emerges is more polarized than when it was travelling in the slab. Thus successive reflections at the first and second surfaces of a series of parallel-sided slabs will produce more and more polarization in the transmitted light. A **pile of plates** (as the arrangement is called) has been used for producing polarized light by transmission (Fig. 759 (ii)), but it must be remembered that without an infinite number of plates the polarization cannot theoretically be complete.

Polarization by Scattering.—We have already mentioned (page 977) that light passing through a collection of small particles is scattered sideways, and we now describe some observations in connection with the polarization of the scattered light. In Fig. 760 ordinary unpolarized light is shown entering the left-hand end of a glass tube containing water in which are suspended fine sulphur particles. The suspension can be made by adding sulphuric acid to a solution of photographic "hypo." If the incident light is white, the scattered light leaving the side of the tube is bluish, because the shorter wave-length components of the white light are more strongly scattered than the red. The light emerging from the other end of the tube is reddish, and this experiment illustrates both the formation of the blue of the sky and the red appearance of the sun at sunset, when the path of sunlight through the atmosphere is longest. If the light scattered at right angles to the incident beam is examined by an analyzer such as the mirror M, it is found to be plane polarized. Thus the strongest reflected beam from the mirror is obtained when the mirror is placed as shown in Fig. 760, *i.e.* with the plane of incidence containing the original beam which is incident on the tube, while if the mirror is rotated about the direction of the scattered beam until the plane of incidence is perpendicular to the original beam, the reflection is a minimum. Therefore, assuming that the mirror reflects only that light which is vibrating parallel to its surface, the vibrations of the scattered light are as shown by the dots in the diagram, *i.e.* perpendicular to the plane containing the incident and scattered beams. This experiment can be used in order to deduce the fact which we have previously quoted, that the plane of vibration is perpendicular to the plane of polarization as we have originally defined it. The argument is that the

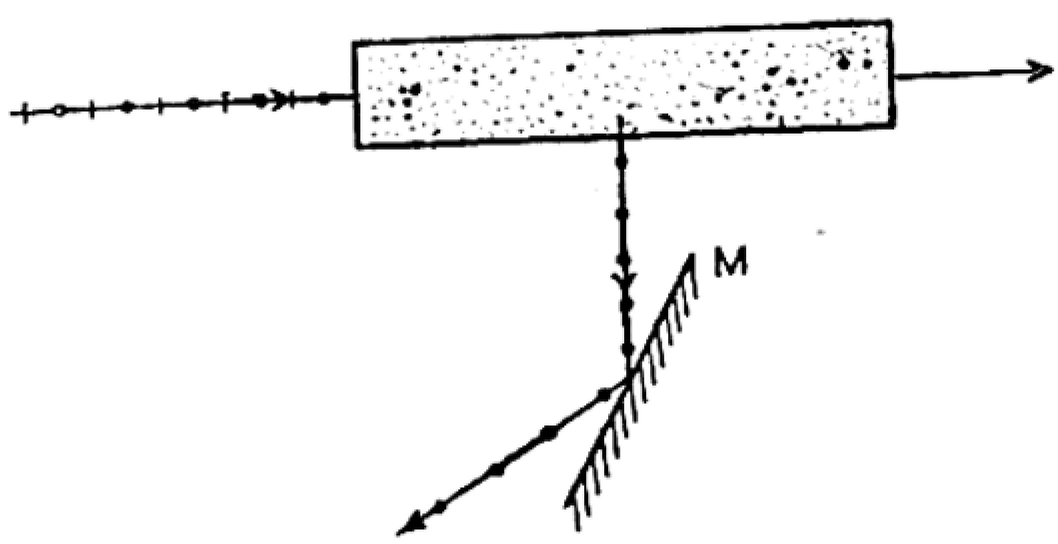


FIG. 760

vibrations of the light scattered through a right angle could not occur in any other direction than perpendicular to the incident beam without there being *longitudinal* vibrations in the incident beam, which, as we have seen, would be incompatible with the phenomenon of polarization.

2. DOUBLE REFRACTION

The Optical Behaviour of a Calcite Crystal.—Calcite (or Iceland spar) is the name given to a naturally occurring colourless crystalline form of calcium carbonate ($CaCO_3$). Its crystals are always in the form of rhombohedra, that is to say they are bounded by six plane faces, each of which is a parallelogram. In each parallelogram one pair of opposite angles is equal to 78° 5′ and the other pair is equal to 101° 55′. Fig. 761 is a drawing of a calcite crystal with the values of some of the angles marked to the nearest degree. The crystal has two blunt corners, A and C, each of which is bounded by three obtuse angles. The **optic axis** is, as we shall see, an important direction to the crystal. Its direction is shown in Fig. 761 as the dotted line AE which is such as to make the same angle with all three edges which meet at a blunt corner. The optic axis is not a single line, however. *Any* line parallel to the direction AE is an optic axis. A **principal section** of the crystal is defined as a plane which is parallel to the optic axis and perpendicular to the crystal face (whether natural or artificial) through which light is entering. A particular principal section is shown in the figure as the shaded plane ABCD, which contains AE and is perpendicular to the small face AB. There are an infinite number of principal sections all parallel to ABCD.

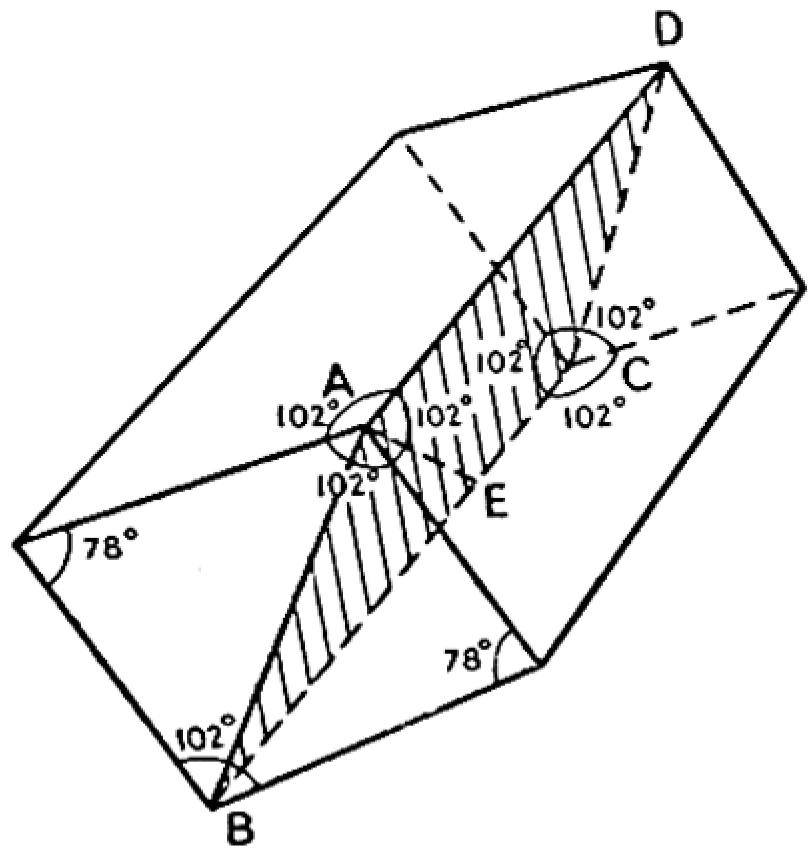


FIG. 761

The optical property exhibited by calcite is called **double refraction** or birefringence, and is made evident in the following observation. A narrow beam of ordinary light is sent into a calcite crystal as shown in Fig. 762 (i). The light is incident normally to one of the faces (which is shown in Fig. 762 (ii)), and the outline ABCD represents a principal section as in Fig. 761. This principal section cuts the face through which the light enters in the line AB (Fig. 762 (ii)). The crystal causes each ray in the incident beam to be split into two separate and distinct rays, known as the **ordinary** and **extraordinary** rays respectively and labelled O and E in the diagram. The refraction of the ordinary ray is of the same type as if the crystal were made of, say, glass. For instance, in

the particular case we are discussing it is undeviated by the crystal, since it passes normally through both the opposite faces. The extraordinary ray, however, does not obey the laws of refraction. It is deviated by the first face although the angle of incidence is zero. An equal and opposite deviation at the second face causes the emergent extraordinary ray to be parallel to but displaced from the ordinary ray. Thus, as might be expected from inspection of the ray diagram, if the crystal is rotated about the incident ray the extraordinary ray moves round the ordinary ray, which remains fixed. In other words, both rays remain in the principal section ABCD. The effects just described can easily be seen if a crystal of calcite is placed over a dot on a sheet of white paper. On looking through the crystal two dots are seen, one produced by the ordinary

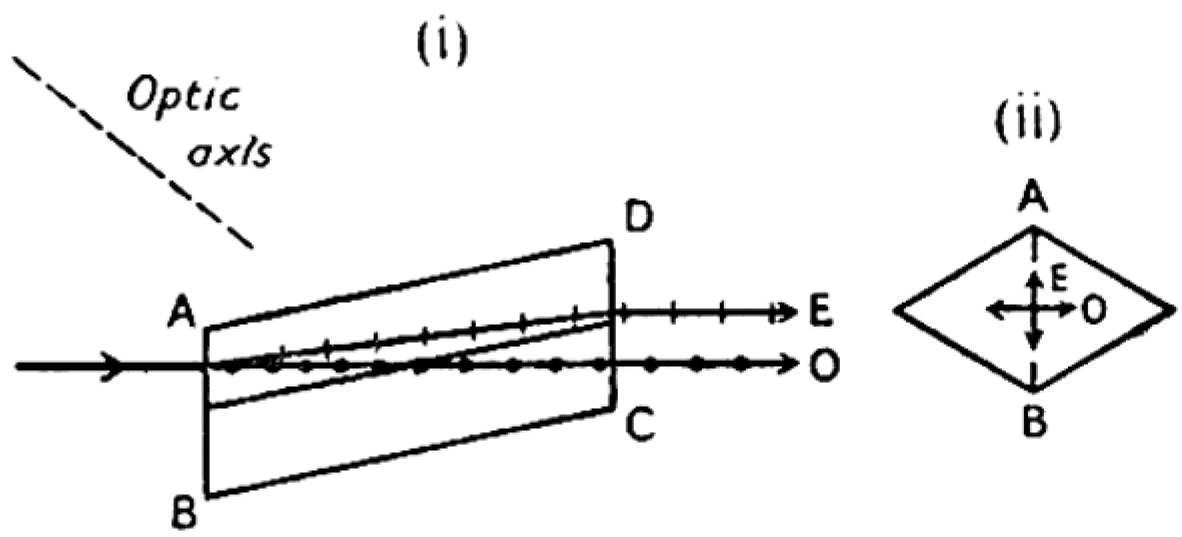


FIG. 762

and the other by the extraordinary ray. When the crystal is rotated on the sheet of paper the image of the dot due to the ordinary ray remains stationary, while that due to the extraordinary ray moves round it.

When the emergent rays are examined for polarization by some sort of analyzer, it is found, as indicated in the diagram, that the direction of the transverse vibration of the ordinary ray is perpendicular to the principal section, and that of the extraordinary ray is parallel to the principal section.

Huygens' Construction for a Uniaxial Crystal.—There are many more observations which can be made concerning the paths of the two rays in calcite in addition to the particular case of normal incidence in a principal section which has just been mentioned. All of these lead to the conclusion that, while the speed of the ordinary wave is independent of its direction of propagation in the crystal, that of the extraordinary wave varies with direction. The two waves are propagated with the same speed only in the direction of the optic axis. Crystals with only one such direction are called **uniaxial.**

If we suppose that a disturbance takes place at a point P (Fig. 763) in a calcite crystal, then P is the origin of two separate and distinct wave-fronts (which are more usually called wave *surfaces* in this connection), representing

respectively the ordinary and extraordinary waves. In the plane containing the optic axis through P, the outlines of the two wave-surfaces are as shown in Fig. 763 (i). The outline of the ordinary wave-surface is a circle because the speed of propagation of the ordinary wave is independent of direction, but for the extraordinary wave the figure is an ellipse which touches the circle at the points A and B, where AB is the direction of the optic axis. Thus, as mentioned above, the two speeds of propagation are equal along the optic axis but are different in all other directions, the extraordinary wave being the faster in calcite. The greatest difference of speed is in the direction perpendicular to

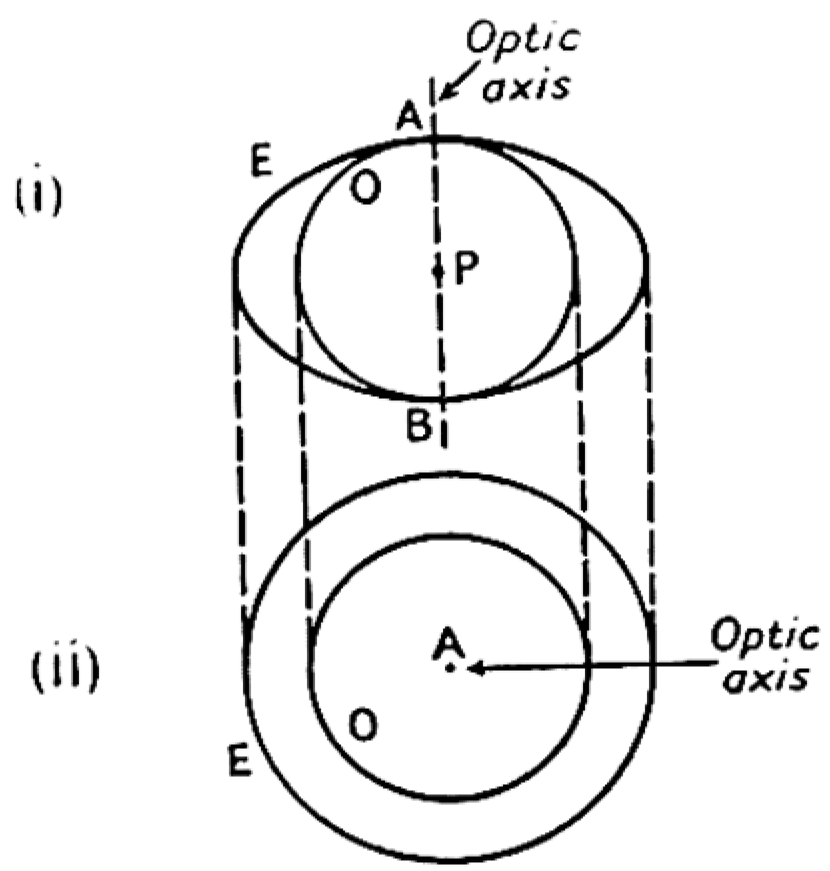


FIG. 763

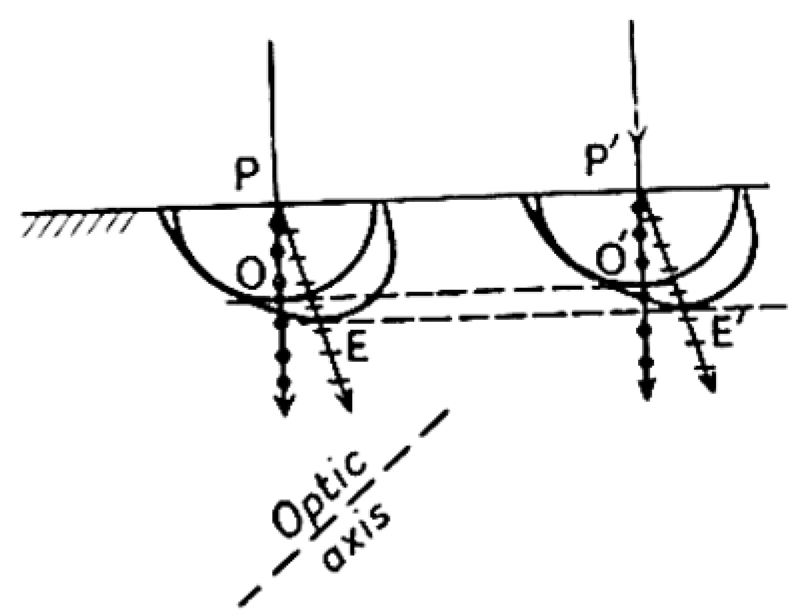


FIG. 764

the optic axis, *i.e.* along the major axis of the ellipse. The complete wave-surfaces are generated by spinning the circle and ellipse in Fig. 763 (i) about the optic axis AB. The circle generates a sphere and the ellipse a spheroid. Fig. 763 (ii) is a section of the two surfaces in a plane through P perpendicular to the optic axis. This plane cuts both surfaces in circles.

The ordinary and extraordinary wave-surfaces can be used in geometrical constructions to illustrate refraction as described on page 695. As a first example, we take the case illustrated in Fig. 762 in which unpolarized light is incident normally on one face of a calcite crystal. Two such rays are incident at points P and P′ respectively in Fig. 764. These two points are regarded as sources of secondary wavelets, according to Huygens' construction, and the diagram shows two identical pairs of ordinary and extraordinary wave-surfaces originating from P and P′ respectively. The optic axis is supposed to lie in the plane of the paper, which, therefore, is a principal section, since the face of the crystal is perpendicular to the paper. The direction of the optic axis decides the relation between the circle and the ellipse, since the line joining the two points at which they touch gives the direction of the optic axis.

In accordance with Huygens' principle, the wave-front of the ordinary wave in the crystal at the instant considered is represented by the line OO′ which is a common tangent to the circular wave-surfaces originating from P and P′. The ordinary ray from P can therefore be regarded as following the path PO, while from P′ it travels along P′O′. Both these directions are obviously normal to the surface. In fact, the ordinary ray is refracted in the same way as light entering a medium which is not doubly refracting. The extraordinary wave-front is represented by the common tangent EE′ to the ellipses, and the paths of the extraordinary rays are PE and PE′. It will be seen that the rays follow paths as already described in Fig. 762 (i). It should be noticed that the speed of propagation of the extraordinary wave-front EE′ in a direction normal to itself is not equal to the rate of increase in length of the extraordinary rays PE and P′E′. Furthermore, the vibration direction of the extraordinary wave is parallel to EE′ and not perpendicular to the rays.

The Nicol Prism.—This is a device which has been frequently used in the past to produce and to analyze plane polarized light. It uses total reflection to eliminate the ordinary ray from the light transmitted by a calcite crystal, thus leaving the extraordinary ray. The principle adopted is to cut a calcite crystal into two parts and to cement them together again with Canada balsam. The total reflection then occurs at the layer of balsam, but it is necessary to arrange for the correct angle of incidence on the layer, and in order to do this, the shape of the calcite crystal is first modified as shown in Fig. 765 (i). A crystal whose length is about three

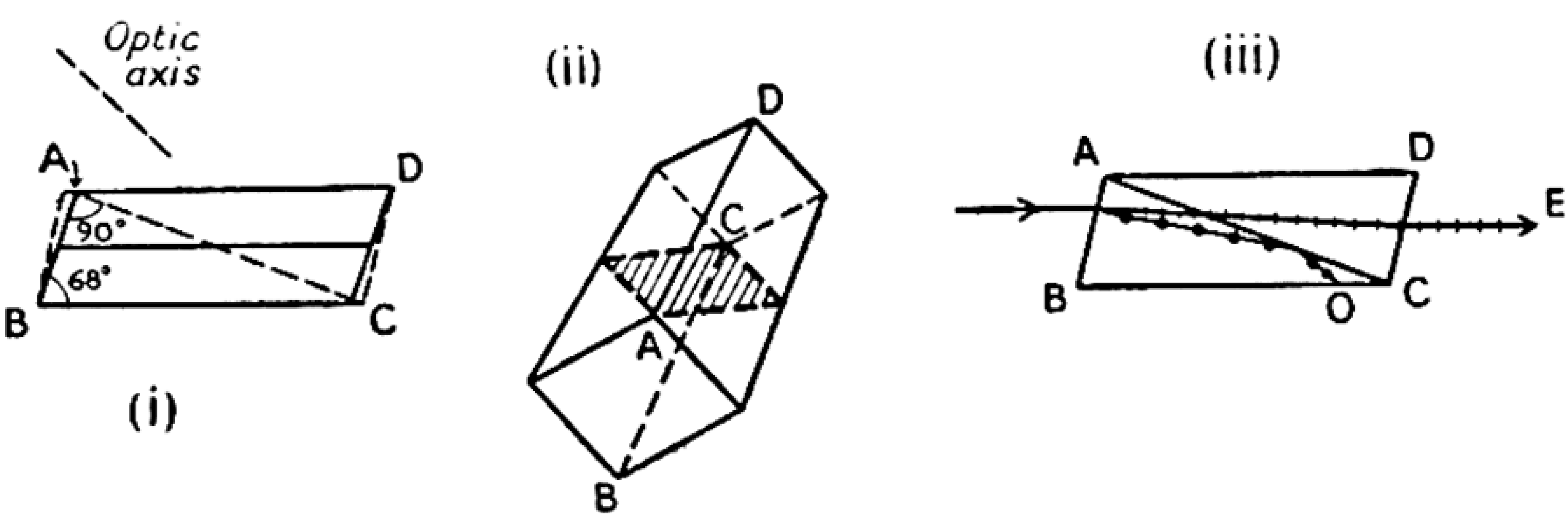


Fig. 765

times its width is chosen, and its small end faces are ground so as to remove the dotted portions indicated in the drawing and to make the angle marked have a value of 68° instead of 71° in the natural crystal. Starting from the blunt corner A, a plane cut is made as indicated by the line AC in Fig. 765 (i) and by the shaded area in the perspective drawing in (ii). It is perpendicular to the principal section ABCD, and is *not* simply made by going from A to C but by cutting at right angles to the face AB. The fact that the cut does actually meet the other blunt corner C is simply due to a correct choice of the length of the crystal to bring this about. The

two halves of the crystal are reunited by a thin layer of Canada balsam placed between them.

The Nicol prism is usually represented diagrammatically as in Fig. 765 (iii). The parallelogram ABCD actually represents the principal section ABCD in (ii), and the diagonal line AC indicates the layer of Canada balsam.

A ray of unpolarized light entering the prism by the face AB in Fig. 765 (iii) gives rise to ordinary and extraordinary rays. On account of the different speeds of propagation (except along the optic axis), the refractive indices of the two rays in calcite differ from each other, that of the ordinary ray being the larger and that of the extraordinary ray varying with direction. The refractive index of Canada balsam lies between that of the ordinary and the appropriate extraordinary values for calcite. It is therefore optically rarer than calcite for the ordinary ray, so that total reflection of this ray occurs if the critical angle (69°) is exceeded. The method of cutting the crystal ensures that this angle is exceeded when light enters the face AB at fairly small angles of incidence. Thus the ordinary ray is totally reflected at the calcite/balsam surface, while the extraordinary ray, although partially reflected, passes through the layer of balsam and out of the opposite face of the crystal. The Nicol prism therefore acts as a polarizer. The plane of vibration of the transmitted polarized light is parallel to the *shorter* diagonal of the end faces (*cf.* Fig. 762 (ii)).

The Nicol prism can, of course, also be used as an analyzer. When unpolarized light is passed through a Nicol prism and then falls on a screen or enters the eye, there is no variation of the intensity of the transmitted light as the Nicol is rotated about the incident beam. The transmitted light has been polarized by the Nicol and the rotation merely alters the plane of vibration. If plane polarized light enters the Nicol, however, the intensity of the transmitted beam is a maximum when the shorter diagonal of the end faces is parallel to the vibration plane of the incident light and zero when the diagonal is perpendicular to the vibration plane. The existence of two positions of zero transmitted intensity for each revolution of the Nicol prism gives evidence that the incident light is plane polarized. When unpolarized light is passed through two Nicols consecutively, the first prism acts as a polarizer and the second as an analyzer. No light will emerge from the second prism if the directions of the shorter diagonals are mutually perpendicular. The Nicols are then said to be **crossed.**

Dichroic Crystals.—There are many other doubly refracting crystals as well as calcite. Some of these substances have the property of absorbing the ordinary ray much more strongly than the extraordinary ray, and are then said to be **dichroic.** Tourmaline is an example, and a crystal of tourmaline will therefore produce plane polarized light (namely the extraordinary ray) when unpolarized light is passed through it. It can

therefore be used as a polarizer or an analyzer in a similar way to a Nicol prism. Two "crossed" tourmaline crystals fail to transmit any light. **Polaroid** is a manufactured film or plate containing dichroic material (*e.g.* herapathite). Among many other uses polaroid is made into spectacles or eye screens for reducing glare in sunlight. One interesting application of such sun-glasses concerns angling. Ordinarily, the light from the sky or direct sunlight reflected by the water surface is strong enough in comparison to light reflected by objects below the surface to mask the latter. Light reflected from the surface is polarized, however, the direction of vibration being parallel to the surface, *i.e.* horizontal. If polaroid spectacles are worn which do not allow horizontal vibrations to pass through them, the reflection from the surface is apparently greatly diminished and objects below the surface are correspondingly more visible.

3. ROTATION OF THE PLANE OF POLARIZATION

Optical Activity.—When monochromatic light is passed through a polarizer and an analyzer there is extinction (*i.e.* no light emerges from the analyzer) when the two are "crossed," because the vibrations of the light from the polarizer are not passed by the analyzer. If, while they are crossed, we interpose a thin plate of quartz cut perpendicular to the optic axis, light emerges from the analyzer. Extinction can be restored by rotating, say, the analyzer through a definite angle θ. This phenomenon can be explained by saying that the quartz has caused a rotation through an angle θ of the plane of polarization (and the plane of vibration) of the plane polarized light which is passing through it. This property of quartz is called **optical activity.** The sense of the rotation can be either left- or right-handed (looking against the direction of the light, *i.e.* towards the source from the analyzer), and according to which direction it is, the quartz is called **lævo-rotatory** and **dextro-rotatory.** The two types of rotation are produced by plates cut from crystals, the shapes of which are mirror images of each other.

The rotation θ for light of a given wave-length is proportional to the thickness of the quartz plate. For sodium light the rotation is 21·7° per mm. thickness of the plate, irrespective of whether it is left- or right-handed. The rotation increases rather rapidly with decreasing wave-length, being about 48° per mm. for the violet light. Therefore, if white light is passed through a polarizer and analyzer with a plate of quartz between them, extinction cannot be obtained in any position of the analyzer, but the colour of the light emerging from it varies as it is rotated.

Many other substances besides quartz are optically active, including solutions. For example, cane sugar is strongly active in aqueous solution. The rotation θ produced by a solution is found to be proportional to the length l of the path of the light in the solution (as in the case with quartz)

and to the concentration c. Thus we can write

$$\theta \propto lc$$
$$= \rho lc$$

where ρ is a constant of the particular substance. It is called the **specific rotation** of the substance when l is expressed in *decimetres* and c in gm. per c.c. of solution. The value of ρ for cane sugar is about 67° per decimetre per unit concentration for sodium light at room temperature.

Polarimetry.—This is the name given to the experimental investigation of the rotation of plane of polarization. The instrument by which it is carried out is called a **polarimeter** or, if it is used mainly for experiments with sugar solution, a **saccharimeter.** We shall describe only the *principle* of the polarimeter without giving details of the various devices used for increasing its sensitivity.

A simple form of polarimeter is shown in Fig. 766. Light from a monochromatic source S, *e.g.* a sodium flame, is made parallel by the converging lens L and then passes through the polarizer P and analyzer A

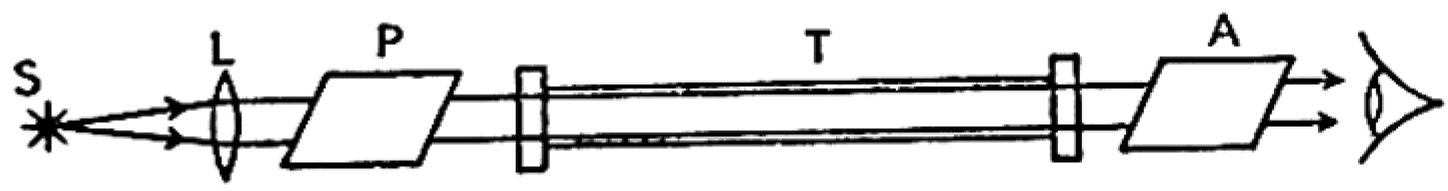


FIG. 766

into the eye. The tube T is often made 20 cm. long ($l = 2$ decimetres), and its ends, which are detachable, are made of optically plane glass. Either P or A can be rotated, and the extent of the rotation can be measured on a degree scale. We shall suppose that A can be rotated. In order to measure the optical rotation due to a particular solution, the tube T is first filled with distilled water and A is rotated until the eye sees no light emerging from A. The reading on the degree scale is taken, and the solution is substituted for the water in the tube T. Light will then be seen to be transmitted by A, which is then turned to restore darkness and the scale reading is again taken. The difference of the two readings gives the rotation of the plane of polarization θ, and if the concentration of the solution is known, the specific rotation ρ can be calculated. On the other hand, since ρ is known for sugar and many other substances, the experiment may be used to determine c, the concentration of any one substance which is known to be present in a solution. This method has been used in medical practice for the estimation of the concentration of sugar in urine, and it is also used in the testing of samples of manufactured sugar.

EXAMPLES LIV

1. How does polarized light differ from ordinary light? Describe how (*a*) to polarize a beam of ordinary light, (*b*) to determine, by a reflection method, the plane of polarization of a beam of plane polarized light.

Either (i) explain the importance of the properties of polarized light to the theory of light, *or* (ii) describe one practical application of plane polarized light. (L.I.)

2. What do you understand by the refractive index of a material?

A Nicol prism is made of calcite, the material being cut through obliquely and cemented together again with a film of Canada balsam. The refractive index of calcite is 1·66 for the ordinary ray and 1·49 for the extraordinary ray in the direction concerned; the refractive index of Canada balsam for both rays is 1·55. Use this information to explain the behaviour of the prism in transmitting only the extraordinary ray, and calculate the least angle of incidence on the film for the ordinary ray to be extinguished.

State briefly how you would verify that the transmitted light is plane polarized. (O.H.S.)

Chapter LV

THE VELOCITY AND NATURE OF LIGHT

1. EVIDENCE OF THE FINITE VELOCITY OF LIGHT

Jupiter's Satellites.—The planet Jupiter, whose orbit round the sun is considerably larger than that of the earth and takes nearly twelve terrestrial years to complete, has several satellites or "moons" revolving round it. As each satellite enters the shadow of the planet it becomes eclipsed, *i.e.* invisible from the earth. The intervals between successive eclipses were observed in 1676 by Römer, a Danish astronomer, in the case of the particular satellite which takes about $42\frac{1}{2}$ hours to complete its orbit round Jupiter. The principles underlying the deductions which he made can be understood from Fig. 767.

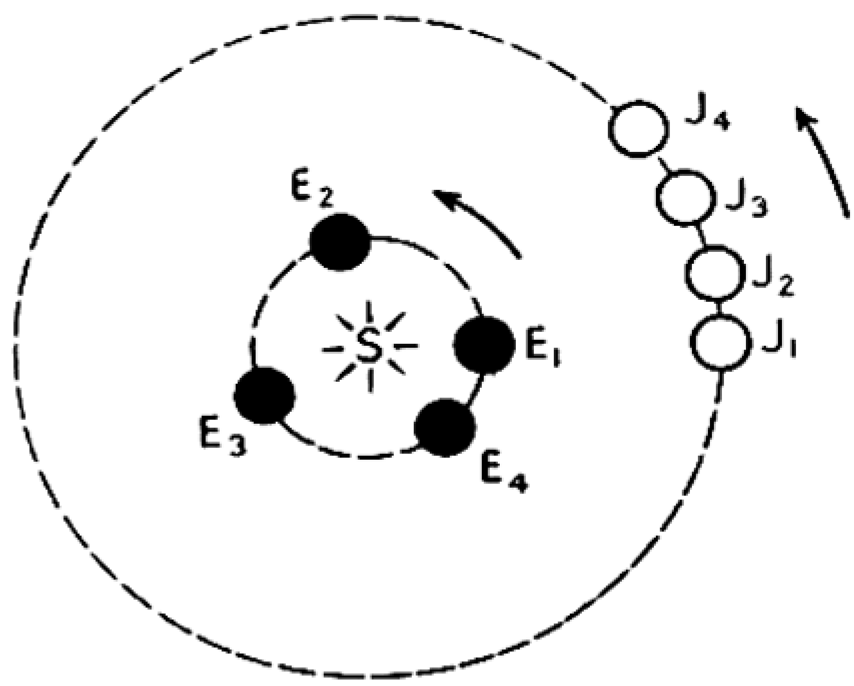


FIG. 767

At a given time the earth and Jupiter are at E_1 and J_1 respectively, and are in line with the sun S. They are then moving parallel to each other. In the later positions E_2 and J_2 their paths are instantaneously mutually perpendicular. They are parallel again at E_3 and J_3 and perpendicular again at E_4 and J_4, and so on.

Suppose that there is a constant interval of time t between the *actual* commencements of consecutive eclipses of one particular satellite, and suppose that light travels with a finite velocity. When the earth and Jupiter are moving parallel to each other the distance between them is almost constant over the time interval t, so that the interval between two successive eclipses will be observed on the earth as t, its actual value. When the earth is moving away from Jupiter, however (as at E_2), the distance between the two planets increases during the interval between two consecutive eclipses, and the light signalling the second of these eclipses has further to travel than that which signalled the first. The *observed* interval between the two therefore exceeds t by the time taken for light to travel the distance by which the separation of the two planets has increased in the time t. This apparent lengthening of the eclipse period occurs for all positions of the earth in the part of its orbit $E_1E_2E_3$, since its distance from Jupiter is increasing all the while it is describing this path. The maximum effect (about 15 sec.)

occurs at E_2, because the rate at which the earth is moving away from Jupiter is greatest at this point. For the next part of the earth's orbit, beginning at E_3, the earth is approaching Jupiter, and the eclipse period is apparently less than t, the greatest discrepancy (15 sec.) occurring at E_4, where the earth is moving directly towards Jupiter, which is at J_4. Thus, as the earth describes its orbit, the observed eclipse period of one of Jupiter's satellites varies continuously, passing through a maximum and a minimum. The true value t could be determined either as an average of all the separate observed values for one complete cycle, or simply as the observed value at such points as E_1 and E_3.

Suppose there are n eclipses between E_1 and E_2. Then, apart from the effect of the finite velocity of light, the eclipse at E_2 should occur at a time nt after the eclipse at E_1. Römer found that the eclipse at E_2 occurred about 11 minutes late according to this method of forecasting. Thus this time represents the sum of all the intervals by which each of the n observed eclipse periods exceeded t. Or, to look at the matter in another way, the 11 minutes is the time taken by the light from Jupiter to travel the distance $(E_2J_2)-(E_1J_1)$, which is approximately equal to the radius of the earth's orbit, *i.e.* 93,000,000 miles. This gives the speed of light as about 140,000 miles sec.^{-1}. It is now known that this is too low an estimate, and that instead of 11 minutes the time is actually about 500 sec., which gives a speed of propagation of 186,000 miles sec.^{-1} or 3×10^{10} cm. sec.^{-1}. Römer interpreted his observations as giving evidence of the finite velocity of light, but this deduction was criticized on the grounds that it assumes that the eclipses actually occur at regular intervals. This is in fact the case, but it cannot be definitely established unless the velocity of light is determined by some independent means.

Bradley's Observations of Aberration.—The effect known as the aberration of light was discovered by the English astronomer Bradley in 1727, and concerns the apparent variation in the position of a star when it is observed from the earth in different positions in the latter's orbit. In principle, it is really a question of relative velocity. Suppose that, as shown in Fig. 768 (i), light from a very distant star reaches the earth E in a direction which is assumed, for simplicity, to be perpendicular to the plane of the earth's orbit. The velocity of the light is shown as c, while that of the earth in a perpendicular direction is indicated by v and is very much smaller than c. The velocity of the light relative to the earth (c') is obtained as shown in Fig. 768 (ii) by subtracting v vectorially from c, as explained on page 13 (Vol. 1). We are concerned with the *direction* of c',

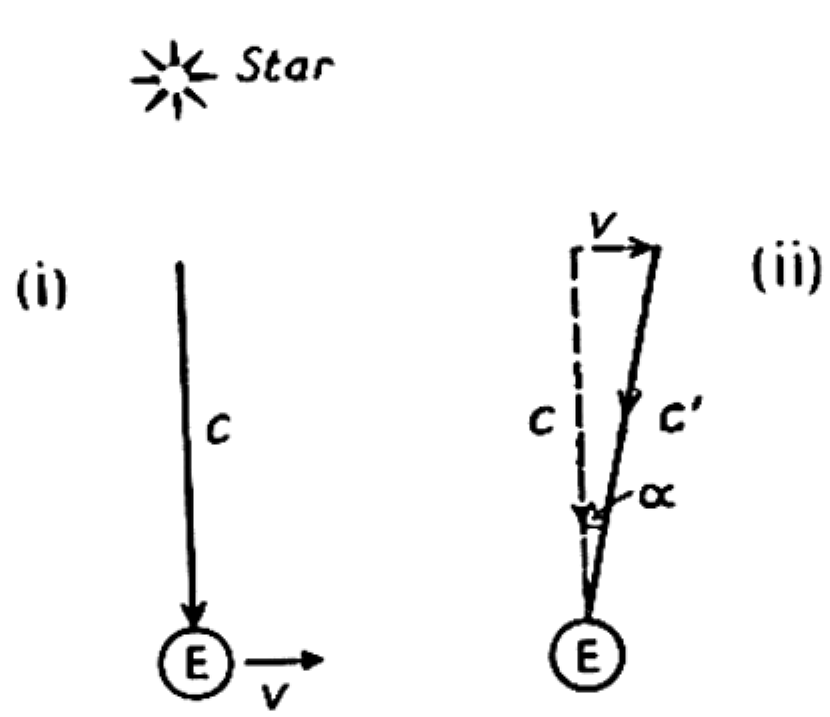


FIG. 768

which is seen to be inclined to that of c by an angle α approximately equal to v/c. This means that if the star is viewed through a telescope, the axis of the instrument must be tilted in the direction of the earth's motion through an angle α from the direction in which the star would be observed if the earth were stationary. The effect is just the same as the apparent slanting of vertical rain which occurs when an observer moves through it. Thus the apparent position of the star is slightly displaced from its actual position in the direction of the motion of the earth. This displacement is not made evident by a single observation, of course, since the actual position of the star is not observable; but as the earth describes its orbit, the direction of the aberration alters continuously, being always in the direction of the earth's motion at the time of observation, and the star therefore appears to describe a small circle (in general an ellipse) every year. This circle is found to subtend an angle of about 41″ at the earth, which means that the angle α is about 20·5″. Since v is known, c can be calculated from this observation, and the result agrees with the value obtained by other methods.

2. TERRESTRIAL DETERMINATIONS OF THE VELOCITY OF LIGHT

Fizeau's Method.—Terrestrial experiments for the determination of the speed of propagation of light are by no means easy on account of the very short time taken by light to travel over observable distances. The first terrestrial experiments to give results were carried out by Fizeau in Paris in 1849. Fig. 769 is a diagrammatic representation of his apparatus.

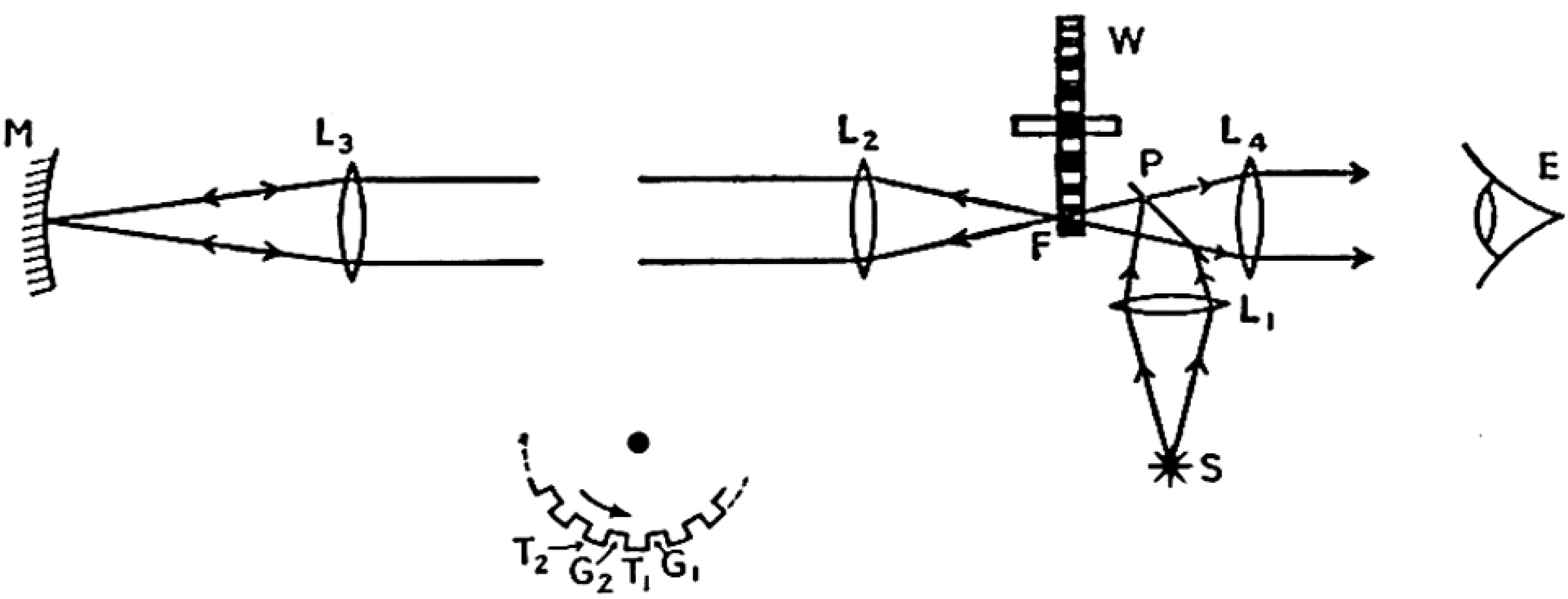


FIG. 769

Light from the source S was converged by the lens L_1 and reflected by the partially reflecting glass plate P so as to come to a focus at F, which was the principal focus of the lens L_2. Parallel light therefore left L_2, and it travelled a distance of nearly $5\frac{1}{2}$ miles until it struck L_3, which brought it to a focus on a concave mirror M, whose radius of curvature was equal to

the focal length of L_3. The light therefore returned to L_3 with a divergence equal to the convergence which it had when travelling from L_3 to M, so that it re-traversed the long distance from L_3 to L_2 as parallel light, and then passed through F. Some of the light therefore returned to S, but some which was transmitted by P passed through an eyepiece L_4 and so entered the eye E of an observer.

A wheel W had a toothed edge which is shown on a larger scale in the drawing on the right. The teeth had the same size as the spaces, and the wheel was mounted as shown in the main drawing so that, as it rotated, the focus F fell alternately on a tooth and a space. Thus the light travelling from L_2 to L_3 was made intermittent by the rotation of the wheel. Suppose that the light passed through the gap G_1 on its outward journey. If the wheel was rotating in the direction shown with a sufficiently high speed, then the tooth T_1 immediately behind the gap G_1 moved forward a certain distance during the time t in which the light travelled from F to M and back. Thus the returning light was partially obstructed by T_1. The next flash was sent out when the gap G_2 crossed the path of the light at F, and T_2 caused a similar obstruction to the returning beam. Thus the intensity of the light seen at E was reduced. When the speed of the wheel was gradually increased, the intensity at E diminished because each tooth moved further into the path of the returning beam in the time t. Eventually a speed was reached at which the intensity was a minimum. This occurred when the rate of rotation was such as to cause a tooth to move into the exact position of the adjacent gap in the time t. After this stage was reached, a further speeding up caused an increase in the light received at E because the next gap began to cross the returning beam, and when the speed was high enough for one gap to move into the exact position of the next in time t the light reached a maximum intensity. The succession of maxima and minima would continue as long as it was possible to increase the speed of the wheel.

In Fizeau's experiment the teeth were the same width as the gaps, and there were 720 of each, so that each occupied $\frac{1}{1440}$ of the rim of the wheel. Fizeau estimated that the first minimum occurred when the wheel was turning at a rate of 12·6 revolutions per sec. At this speed the time t for a tooth to move to the position of the adjacent gap was therefore $\dfrac{1}{12{\cdot}6 \times 1440}$ sec. The distance from F to M was 8633 metres, so that the total distance travelled by the light in time t was 2×8633 metres. This gives the velocity of light as $(12{\cdot}6 \times 1440 \times 2 \times 8633)$, or about $3{\cdot}14 \times 10^8$ metres sec.$^{-1}$.

Fizeau's experiment is important historically, but it was not very satisfactory from the point of view of accuracy. Obviously it would not be easy to decide when minimum intensity had been reached. The light reaching the observer's eye was weak in any case, and the difficulty of judging the minimum was further increased by the general illumination of

the field of view by light which was reflected back by the teeth of the wheel at F. Subsequent workers, notably Cornu and Young and Forbes, improved the original Fizeau experiment, but in more modern times the rotating-mirror method which we are about to describe has proved more successful than toothed-wheel methods.

Foucault's Method.—The use of a rotating mirror is said to have been suggested by Wheatstone and by Arago, but the experiment was first performed by Foucault in France in 1850. The principle of the method is shown in Fig. 770. O was a strongly illuminated object, which in Foucault's experiment consisted of five lines scratched on silvered

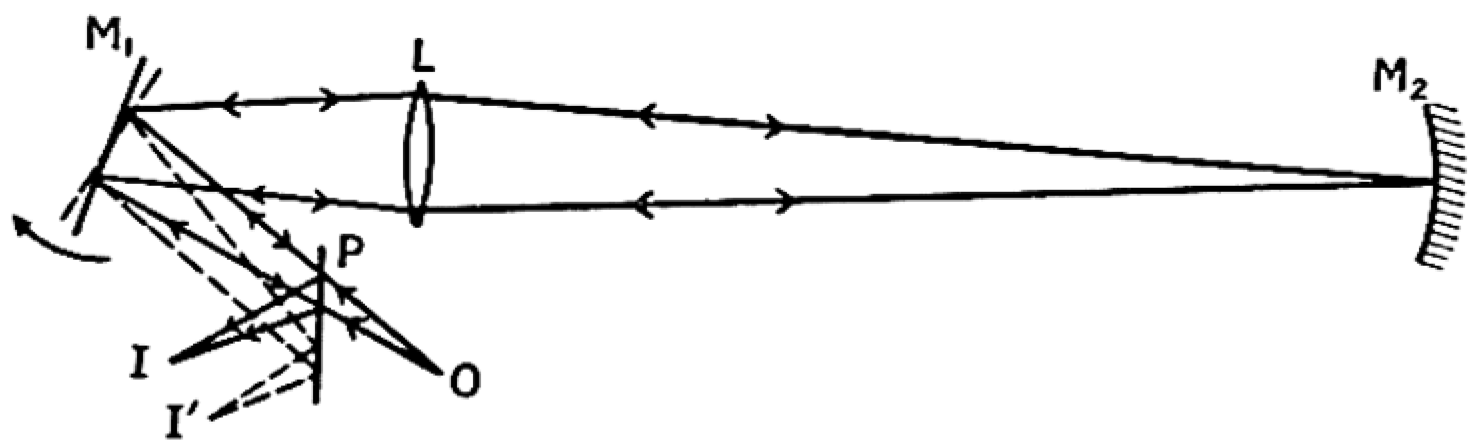


FIG. 770

glass through which strong sunlight was sent by means of a mirror (which is not shown in the drawing). The light from O passed through a partially reflecting plate P, was reflected by the plane mirror M_1 and then brought by the converging lens L to a focus on the surface of the concave mirror M_2, the radius of curvature of which was equal to the distance LM_2. As a result of the reflection at M_2 the light retraced its path, and that portion of the returning light which was reflected by P formed an image of O at I. The mirror M_1 could be rotated about an axis perpendicular to the plane of the drawing. If the speed of rotation was high enough, then during the short time taken by the light to travel from M_1 to M_2 and back, M_1 turned through a sufficiently large angle (as indicated by the dotted line) to cause the returning light to follow the dotted path from M_1 to P and thus to produce a shift of the image from I through a measurable distance to I′. This displacement corresponds to a change in the direction of the returning light through an angle of $\frac{II'}{(IP)+(PM_1)}$, *i.e.* $\frac{II'}{OM_1}$ radians. Therefore the mirror M_1 has turned through half this angle in the time taken for the light to travel from M_1 to M_2 and back again. If the steady speed of rotation of the mirror is observed to be n revolutions per second, then it describes $2\pi n$ radians per second, and the time in which it turns through the angle $\frac{II'}{2(OM_1)}$ is $\frac{II'}{4\pi n(OM_1)}$ sec. The light has travelled a distance $2(M_1M_2)$ in this time,

so that its speed of propagation is

$$\frac{8\pi n(OM_1)(M_1M_2)}{II'}$$

In Foucault's original experiment the focal length of the lens L was about two metres. The light was not, in fact, reflected immediately back to L from M_2, but was sent by M_2 on to another similar mirror and thence to a third, and so on, and it was not until the fifth mirror had been reached and the light had travelled about 20 metres from M_1 that it was reflected back along the same path by the last mirror.

Foucault's rotating mirror was driven by an air turbine. In order to maintain its speed constant and to determine this speed, a toothed wheel was driven at a constant known speed by a clockwork mechanism and was viewed in the light reflected by the rotating mirror which, of course, consisted of a series of intermittent flashes at any one place. The speed of the mirror was then adjusted until the wheel appeared stationary, and it was then known that the mirror made a complete rotation in the time during which one tooth moved into the position of the next. This time could be calculated from the speed of the wheel and the number of teeth round its circumference. The intermittent nature of the light reaching the displaced image I′ is, of course, not evident to the eye because of the rapidity of the flashes, but the intensity of I′ is much smaller than that of I, because the flash lasts for only that fraction of a complete revolution of M_1 for which the light falls on M_2. This was a serious drawback in Foucault's experiment, as well as the fact that the displacement II′ was less than one millimetre.

We have seen (page 696) that the refraction of light can be accounted for on the wave theory of light if we suppose that the speed of propagation in a medium of refractive index n is equal to $\frac{c}{n}$, where c is the speed *in vacuo*. That is to say, the speed in any other medium is less than in a vacuum. Foucault tested the truth of this statement by using another concave mirror M_2' (Fig. 771) in addition to M_2, and placing a long tube T filled with water between M_1 and M_2'. As M_1 rotated, one displaced image was produced by the light travelling in air between M_1 and M_2, and another was produced by light which travelled the same distance through T. Foucault was able to detect that the latter image was the more displaced, thus showing that the light had travelled more slowly in the water. As a result of a similar but much improved experiment, Michelson showed in 1885 that the ratio of the velocity of light *in vacuo* to that in

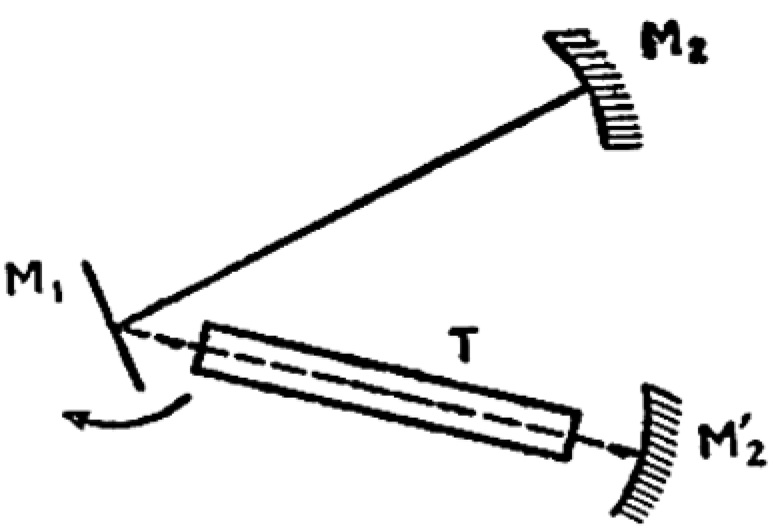


FIG. 771

water was, within the limits of his experimental error, equal to the refractive index of water (1·33).

Michelson's Determinations.—We shall now describe the principle of Michelson's determination of the velocity of light in air which was made in U.S.A. in 1926 and was one of the most accurate determinations yet made. The arrangement used is shown diagrammatically in Fig. 772. Light from a powerful arc lamp passed through the slit S and then fell on M_1, which consisted of a rotating eight-sided drum, each side of which was a plane mirror. When any one of the mirrors passed through the correct position, the light was reflected by the mirror m_1 on to m_2 which

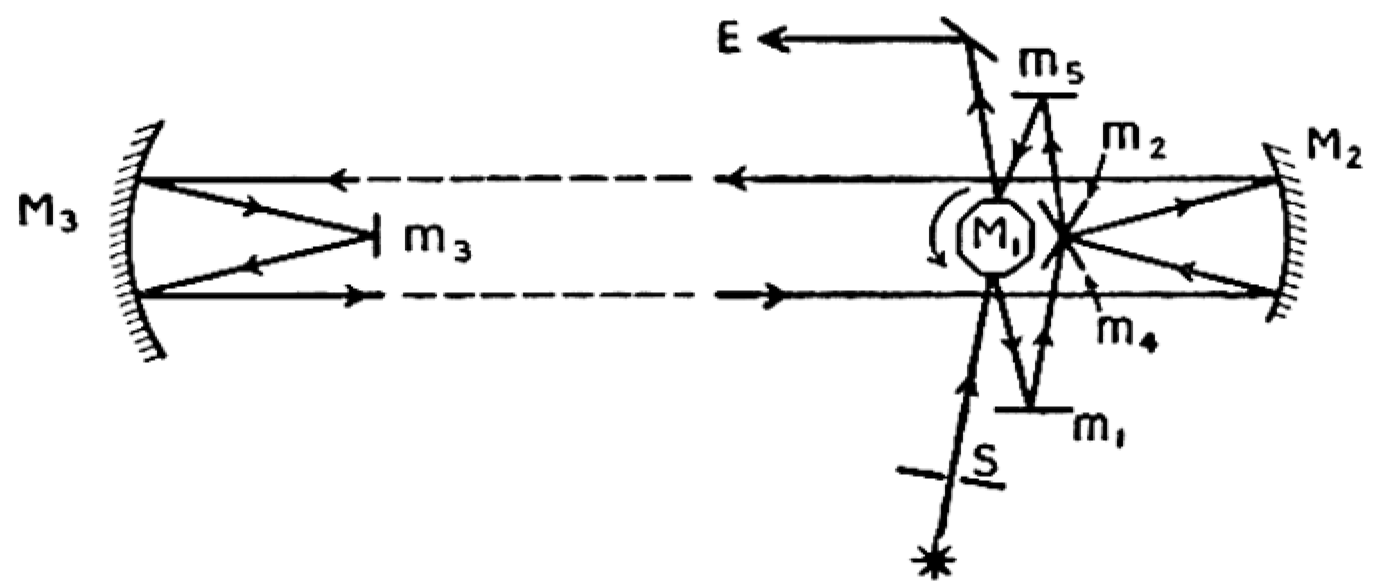


FIG. 772

was at the principal focus of the large concave mirror M_2. The part of the apparatus so far described was set up at Mount Wilson Observatory. The light was reflected by M_2 to a similar mirror M_3 situated on a mountain 22 miles away. Here it was reflected to m_3 at the focus of M_3 and so back to M_3 which returned it to M_2, whence its path was m_4, m_5, M_1, and so to an eyepiece at E in which an image of the original slit S was formed. If the drum of mirrors M_1 were rotated at such a speed that during the time taken by the light to traverse the path from M_1 and back again each of the eight mirrors on M_1 moved into the position previously occupied by its neighbour (*i.e.* a rotation of the drum of 45°), then the returning light was reflected just as if M_1 had been stationary, and the image in E was not displaced by the rotation. This condition occurred in Michelson's experiment when M_1 was rotating at 528 revolutions per second. The time taken for the rotation of 45° was therefore $\dfrac{1}{8 \times 528}$ or 0·00024 sec., and when the carefully measured distance traversed by the light was divided by this time, the value of the velocity of light was obtained. The experiments gave an average value of 299,796 km. sec.$^{-1}$ for the velocity of light *in vacuo*.

During the actual determination the light was, of course, travelling through the atmosphere, and the speed *in vacuo* was calculated by multiplying the observed value by the refractive index of the atmosphere under prevailing conditions. There was naturally some uncertainty with

regard to the value of the refractive index of a column of the atmosphere 22 miles long, and so Michelson planned a direct determination of the speed *in vacuo* which was carried out in 1931 after his death. The principle of the method was the same as the Mount Wilson experiment in that a rotating drum of mirrors was used. The light travelled in an evacuated iron pipe about a mile long. It was reflected to and fro in this pipe a sufficient number of times to give it an effective path of about 10 miles between leaving and returning to the rotating mirror. The mean value obtained was 299,774 km. $sec.^{-1}$.

Kerr-Cell Methods.—Owing to the fact that the rate at which a beam of light can be interrupted by such devices as the toothed wheel and the rotating mirror is limited by mechanical considerations, the light must necessarily travel a long distance in experiments embodying either of these methods. For instance, even with the path of 44 miles in the Mount Wilson determination, the mirror drum had to rotate at over 500 revolutions per second, giving a rate of interruption of the order of 4000 per second.

A much faster rate of interruption of a beam of light can be achieved by a non-mechanical method using a device known as a Kerr cell. This depends upon the phenomenon called the Kerr effect after its discoverer. If two flat conducting electrodes are placed in a liquid such as nitrobenzene, the liquid immediately becomes doubly refracting when an electric potential difference is set up between the electrodes, and the effect disappears at the instant when the electric field is removed. As we have seen in the previous chapter on Polarization, when a polarizer and an analyzer (*e.g.* Nicol prisms) are "crossed," the plane polarized light emerging from the polarizer fails to get through the analyzer because its plane of vibration is perpendicular to the plane in the analyzer in which vibrations are transmitted. When a doubly refracting material is placed between the polarizer and analyzer, its effect in general is to produce a component of vibration in the direction which the analyzer transmits, so that some light emerges from the latter. Evidently, therefore, if a Kerr cell is placed between a polarizer and analyzer which are crossed, the arrangement will transmit light when the potential difference exists, but not otherwise, because the nitrobenzene is not normally doubly refracting. An alternating or interrupted potential difference of frequency many millions of cycles per second can easily be generated by means of a valve oscillator, and when this is applied to the electrodes of the Kerr cell, light interrupted at this frequency emerges from the analyzer.

Early determinations of the velocity of light with the help of Kerr cells were modelled on the rotating-mirror determinations in that two separate cells were operated at the same frequency. Light passing from one cell to another at a known distance from it, would either pass through the second cell or be cut off by it according to whether the light took an even or an odd number of half-periods of oscillation of the potential difference to travel the distance between the two cells.

A more successful method of determination used by Anderson in 1941, however, employed only one cell. Fig. 773 indicates only the *principle* of the method. After passing through the Kerr cell K, the now interrupted light was divided into two parts by a partial reflector P. Each part was then sent back to P by the mirrors M_1 and M_2 respectively, and thence to a photo-cell detector C, the two paths being of different lengths. The photo-cell was connected to a circuit designed to give a strong response when light pulses were arriving with a frequency equal to that of the potential difference applied to the Kerr cell and no response when the light falling on the photo-cell was steady. The latter condition would be fulfilled if the difference of path were equal to the distance travelled by the light pulses in an odd number of half-periods of the oscillating potential difference, because in this case the pulses in one beam would arrive exactly between the pulses of the other and the effect would be that of an uninterrupted beam. The path differences satisfying this condition could be approximately calculated using the known frequency of the potential difference and the value of the velocity of light determined by previous methods. The mirrors were so placed as to give an anticipated path difference corresponding to $11\frac{1}{2}$ periods, and were then further adjusted until there was no response in the photo-cell circuit. Then the path difference was measured accurately and was known to be the distance travelled by light in $11\frac{1}{2}$ periods. Hence the velocity could be calculated and a mean value of 299,776 km. sec.$^{-1}$ was obtained.

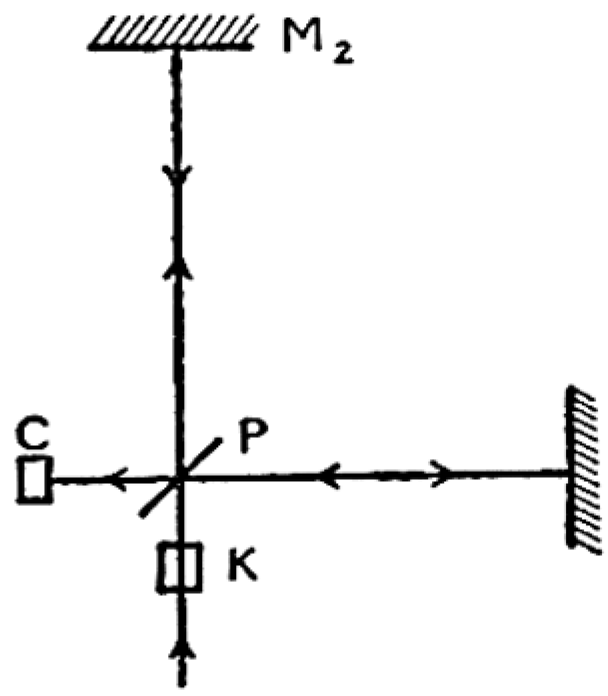


FIG. 773

Group and Phase Velocity.*—It will possibly come as a surprise to the student to learn that experiments to determine the velocity of light do ***not*** in fact measure the speed of the individual light waves in a light beam. It is evident that in terrestrial experiments (in all of which interrupted light is used), as well as in the observations of Jupiter's satellites, what actually is observed is the velocity of the beginning and end of a wave-train, *i.e.* of a collection or group of waves such as is indicated in Fig. 774. If uninterrupted light waves of constant amplitude are travelling from a source to the eye or some other detector, we have no means of observing the speed of propagation of the individual waves or the **phase velocity** as it is called. All we can do is to measure the speed of some superimposed effect such as a change of amplitude or a chopping of the waves into groups as in the experiments mentioned. It can be shown theoretically that if the medium in which a group of waves is travelling is dispersive, *i.e.* if the speed of propagation

FIG. 774

* This section is more advanced than the general standard of this book.

varies with wave-length, then the group and phase velocities are different. For example, the group of surface waves which moves outwards from a place where a stone has been thrown into a pond travels more slowly than the crests and troughs of the individual waves within it. In other words, the waves continually move through the group from rear to front, dying out when they reach the front and being created at the rear.

The full theory of this phenomenon is much too advanced for a book of this standard, but the underlying principle can be briefly explained. It is mentioned several times in Vol. 3 that a periodic disturbance which is not purely sinusoidal, such as a sound wave from a musical instrument, can be represented as the resultant of a Fourier series of sinusoidal waves having frequencies which are proportional to the whole numbers. A *single* disturbance like a group of waves can also be shown to be the equivalent of a series of sinusoidal waves, but in this case their wave-lengths differ from each other by infinitesimal amounts, *i.e.* all possible wave-lengths can be present. The character of the disturbance, *e.g.* whether it is a short wave-train like that shown in Fig. 774 or whether it is a single pulse, decides the relative amplitudes of the sine waves, *i.e.* which wave-lengths are predominant and which are relatively unimportant. The sum of the component waves is called a **Fourier integral,** as distinct from a Fourier series.

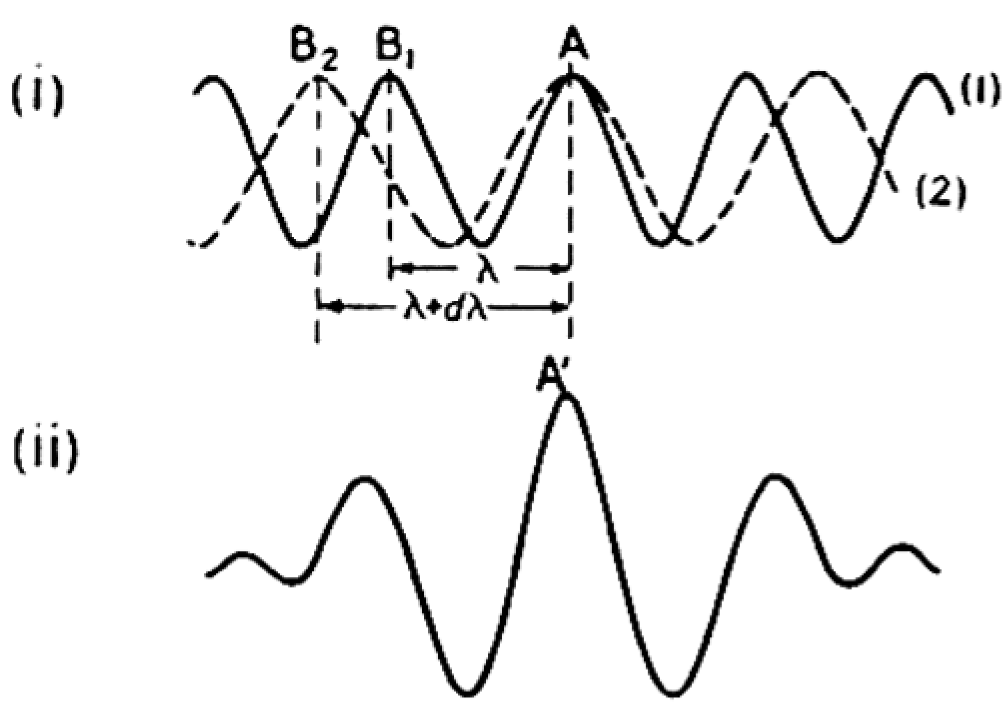


FIG. 775

In order to make possible a simple discussion, it is usual to consider a group which is the equivalent of only two Fourier components. This is illustrated in Fig. 775, in which (i) shows the two individual sine waves ((1) and (2)) of equal amplitude but slightly different wave-lengths, λ and $\lambda + d\lambda$. The resultant of the two components is shown in (ii), and it will be seen that this represents what we have called a "group" in the sense that it is something which can be detected on account of the variation of amplitude like the groups into which the light beam is chopped in a toothed-wheel or rotating-mirror experiment. Fig. 775 (ii) is, so to speak, a rudimentary group of the form shown in Fig. 774. Suppose that the waves (1) and (2) are travelling to the right with velocities v and $v + dv$ respectively. At the instant considered, two crests are coinciding at A, and together they form the maximum amplitude of the group at A′. We now relate the movement of A′ to that of the two individual waves. Immediately after the instant depicted in Fig. 775 wave (2) begins to get ahead of wave (1), so that the two crests no longer coincide at A. After a finite interval the two crests B_1 and B_2 coincide and produce the maximum

displacement in the group. Thus the group moves in the same way as does the point of coincidence of waves (1) and (2). The interval between coincidence at A and coincidence at B_1 is the time necessary for B_2 to overtake B_1. Since $B_1B_2 = d\lambda$, and the velocity of wave (2) relative to wave (1) is dv, this time is $\frac{d\lambda}{dv}$. In this time the maximum of the group, and in fact the whole group, has moved to the left relative to wave (1) through a distance AB_1, which is equal to λ, so that the velocity of the group to the left relative to wave (1) is $\lambda \div \left(\frac{d\lambda}{dv}\right)$, or $\lambda\frac{dv}{d\lambda}$. The velocity of wave (1) is v to the right, so that the actual velocity of the group is $\left(v - \lambda\frac{dv}{d\lambda}\right)$. Calling the group velocity u, therefore, we have the equation

$$u = v - \lambda\frac{dv}{d\lambda} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Although this is the equation which is usually said to relate phase and group velocities, it should be realized that we have established it for a certain kind of group which can be represented by only two separate sine waves. In general, the representation of any given group as a collection of sine waves requires a large number of such waves extending over a finite wave-length range. In such a case u varies with λ, so that different parts of the group have different velocities. In other words, the group changes its form as it travels.

As already stated, terrestrial determinations of the velocity of light really measure the velocity of groups of light waves, *i.e.* u and not v. At one time it was thought that Bradley's observations of aberration measured the phase velocity v, but it has since been shown that this is not the case. The fact is that the *energy* of the light waves can be shown to travel with the group velocity, and it is the arrival of energy at the detector which is observed, even in the aberration experiment. In any case, free space is considered to be a non-dispersive medium $\left(\frac{dv}{d\lambda} = 0\right)$, so that u and v have the same values in it.

The difference between u and v becomes important in connection with direct determinations of the velocity of light in liquids such as were made by Michelson. Thus Michelson found the ratio of the velocity in space to that in carbon bisulphide to be 1·76, whereas the refractive index of this liquid is known to be only 1·64. The reason for the difference lies in the fact that the refractive index of a medium as directly determined using Snell's law represents the ratio of the *phase* velocities in free space and in the medium, whereas the direct determinations of velocities give the *group* velocities. Thus if the velocity *in vacuo* (both phase and group) is c, then the phase velocity (v) in carbon bisulphide is $c/1{\cdot}64$,

while the group velocity (u) is $c/1{\cdot}76$. The value of $\dfrac{dv}{d\lambda}$ for carbon bisulphide can be found from measurements of refractive index for different wave-lengths, so that u can be calculated from v by equation (1), and it is found to agree with the value $c/1{\cdot}76$.

The Doppler Effect.—This phenomenon, which concerns the effect of the motion of the source and observer on the apparent frequency of the radiation emitted by the source, has been discussed in connection with sound waves (page 577, Vol. 3). It is, of course, very noticeable in sound, *e.g.* with locomotive whistles and car horns; but with light the very much higher speed of propagation so reduces the order of magnitude of the effect as to make it appreciable only in special circumstances. The effect of the motion of the source has been worked out on page 578. It can be set out briefly as follows. Suppose that a light source emitting at a frequency ν is travelling towards an observer with a velocity u. If c is the speed of propagation of the light, then in the unit time in which the light travels a distance c from the source towards the observer, the source has travelled a distance u in the same direction. Therefore the light waves which, if the source were stationary, would have occupied a distance c, are, so to speak, compressed into a distance $(c-u)$. The length of each wave is therefore $\left(\dfrac{c-u}{c}\right)$ of its normal length, which is $\dfrac{c}{\nu}$. Therefore the wave-length is $\dfrac{c-u}{\nu}$, and the frequency ν' which is apparent to the observer is the quotient of c by this wave-length, so that

$$\nu' = \nu \frac{c}{c-u} \qquad . \quad . \quad . \quad . \quad . \quad (2)$$

If the velocity of the source u is directed away from the observer, the above argument gives the relation

$$\nu' = \nu \frac{c}{c+u}$$

When the source is stationary and the observer is moving towards it with a velocity v, the velocity of the observer relative to the waves is $(c+v)$ instead of c, which it would be if the observer were stationary. Therefore the number of waves which pass the observer in unit time, *i.e.* the apparent frequency ν', is $\dfrac{c+v}{c}$ of what a stationary observer would experience. Therefore

$$\nu' = \nu \frac{c+v}{c} \qquad . \quad . \quad . \quad . \quad . \quad (3)$$

For an observer moving away from the source the relation is

$$\nu' = \nu \frac{c - v}{c}$$

In discussing the Doppler effect in sound it is emphasized that the relations are different according to whether the source or observer (or both) are moving. That is to say, the apparent frequency is not determined by the *relative* velocity of source and observer but by their separate velocities through the medium in which the sound was propagated. Now the medium in which light is propagated (if it is taken for granted that a medium is necessary) is the hypothetical "ether," and it might well be assumed that the earth is in motion relative to the ether in view of the known motion of the earth with respect to the sun and of the sun with respect to other stars. The famous Michelson-Morley experiment was designed to detect this motion by comparing the velocity of light on the earth in two mutually perpendicular directions. No motion of the earth relative to the ether was detected, and this negative result subsequently formed the basis of the **theory of relativity,** one of the postulates of which is that the detection of such a motion is a fundamental impossibility. If this is so, then it is meaningless to ascribe separate velocities to the source and observer—only their relative velocity has any significance. Accordingly, the formula for the Doppler effect, as calculated from the theory of relativity, differs from equations (2) and (3). However, both equations (2) and (3) and the corresponding relativity equation become identical in cases where the square of the ratio of the relative velocity to c is negligible, and this condition is very frequently fulfilled.

The frequency change due to the Doppler effect is detected by the change in the position of identifiable spectral lines in the spectrum of the source. The velocities of stars relative to the earth have been measured in this way. If the spectral lines due to the elements present in a star are shifted towards the red end of the spectrum, so that their wave-lengths appear to be longer than normal (*i.e.* their frequencies are lower than normal), then the star is known to be receding from us. An approaching star gives a shift in the reverse direction. The Doppler effect has also been used to detect the rotation of the sun, there being an apparent difference in wave-length between the spectra of the radiation from opposite edges of its disc. The rotation of Saturn's rings and of double stars has been detected in the same way. Direct observations of the Doppler effect have been made in terrestrial experiments involving rapidly moving mirrors.

3. THE NATURE OF LIGHT

The Wave Theory of Light.—In discussing the interference, diffraction and polarization of light in the foregoing chapters, we have almost taken it for granted that light is a wave motion. It is true that the wave nature of

light is rather lost sight of in the chapters on lenses and mirrors, in which the conception of the transmission of light energy along rays is used, but it has been explained that this geometrical treatment is not inconsistent with the wave theory of light provided we are not concerned with such comparatively small-scale effects as the diffraction of light round the edge of an obstacle. It can be said, therefore, that the wave theory of light adequately explains the phenomena we have dealt with in the preceding chapters.

The name of Huygens is particularly connected with the early development of the wave theory. Near the end of the seventeenth century he gave his explanation of reflection and refraction in terms of it, but he was unable to account for polarization because he was thinking of light as a longitudinal motion. It was not until over a hundred years later, *i.e.* during the first part of the nineteenth century, that Young suggested that polarization would be explained on the assumption that light waves are transverse. Fresnel developed this idea and introduced his theory of half-period zones to account for rectilinear propagation.

The Corpuscular Theory of Light.—Newton, who was a contemporary of Huygens, did not accept the latter's form of wave theory, and his own great reputation gained many supporters for a corpuscular theory of light, which regarded light as a stream of particles moving in straight lines. One of Newton's objections to the wave theory was, naturally enough, the common observation that waves in fluids bend round obstacles and do not produce sharp shadows. It was eventually shown, as we have seen, by Fresnel's theory of zones that the apparent difference between the behaviour of light and of, say, water waves is simply due to the comparatively short length of light waves. Until this was realized, however, the corpuscular theory appeared to commend itself by the ease with which it explained rectilinear propagation as an example of Newton's first law of motion. It is true that Newton was aware of the diffraction of light at the edge of an obstacle (he called it "inflexion"), but he explained this phenomenon by supposing that the obstacle exerted forces on the light corpuscles as they moved past its edge.

Reflection appeared to be readily explained on the corpuscular theory by supposing that the light particles were repelled by the reflecting surface. The change of direction of a ray when it is refracted into an optically denser medium was explained by an *attraction* exerted on the particles by that medium. In order to account for the possibility of reflection and refraction taking place simultaneously, Newton supposed that during their motion through a medium the particles were put (perhaps by waves which accompanied them) alternately into phases or "fits" of easy reflection and easy transmission. This part of his theory was suggested to Newton by his observations of what we now call Newton's rings, and it is interesting to note that the agent responsible for the alteration in the properties of the particles was thought of by Newton as a *wave* motion

which was generated by the interaction of the particles and matter and travelled faster than the particles themselves.

The corpuscular theory survived into the first part of the nineteenth century in spite of the satisfactory way in which the wave theory was shown to explain interference and diffraction by Young, Fresnel and others. What was generally regarded at the time as a final piece of conclusive evidence against the corpuscular theory came in connection with the refraction of light. We have already mentioned that the theory explained the bending of light towards the normal by supposing that the light corpuscles were attracted to the optically denser medium. This would cause them to travel *faster* in the denser medium than in free space, whereas the wave-theory explanation of the same phenomenon required the light waves to travel *more slowly* in this medium. It was to test this point that a direct determination of the velocity of light in water was made by Foucault in 1850. The result showed that the velocity was lower in water, although it was not until thirty years later that Michelson proved that the ratio of the velocities in air and water was equal to the refractive index of water, as required by the wave theory.

The Dual Nature of Light.—The somewhat difficult postulate of an all-pervading very dense and rigid ether which appeared at first to be necessary for the transmission of the transverse light waves at high velocity was no longer required by the wave theory after Maxwell had introduced his theory of electromagnetic radiation, which was supported experimentally by the discovery of radio waves. At the end of the nineteenth century, therefore, the wave theory appeared to be firmly established, but during the first part of the twentieth century it began to show signs of being unable to account for all the known phenomena.

We have seen (page 543, Vol. 2) how, in order to account for the distribution of energy in the spectrum of a black body, it was necessary to replace the classical idea of continuous emission by the **quantum theory,** which supposed that the smallest amount of radiant energy that could be emitted by a source was not infinitesimal but was equal to $h\nu$, where ν is the frequency of the radiation and h is Planck's universal constant. This was only one example of the necessity for postulating a kind of atomicity in radiation which was contrary to the classical wave theory.

Another remarkable instance occurred in connection with the **photo-electric effect,** which is the name given to the emission of electrons from a metal surface when light (or other radiation) falls on it. In order that an electron shall just be able to pass outwards through the surface of a piece of metal, it must acquire from the radiation a certain minimum amount of energy p, which varies from metal to metal. Whatever energy the radiation gives the electron in excess of p appears as kinetic energy of the electron when it is free from the surface. Suppose that electrons are being ejected from a particular metal surface by the action of monochromatic light falling uniformly on the latter. If light were a wave

motion, we should presume that if its intensity were reduced, the light energy available for each electron upon which it fell would be correspondingly diminished, with the result that each would be ejected with less kinetic energy than before, although the number of electrons emitted in a given time would be unaltered. Further continuous reduction in intensity would, we should suppose, eventually reduce the energy available for each electron below the value p and the photo-electric emission would then abruptly cease. This is *not* found to be so in practice. Reducing the intensity of the light has no effect on the energy of the individual electrons emitted, but it diminishes the number of electrons set free in a given time. Einstein pointed out that these observations could be explained by attributing a particle rather than a wave character to light. Thus, if the incident beam of monochromatic light of frequency ν is regarded as a stream of particles, quanta or, as they are now called, **photons,** each of energy $h\nu$, then each individual electron in the metal surface can either acquire the whole energy of a photon or none at all. Each electron which absorbs a quantum of energy would use an amount p in getting free from the surface, and would therefore be ejected with a kinetic energy equal to $(h\nu - p)$. Reducing the intensity of the light would mean diminishing the number of photons falling on the surface, thereby diminishing the number of electrons emitted but not their individual energy.

The relation put forward by Einstein, namely,

$$E = h\nu - p$$

where E is the kinetic energy of the escaping electron, shows how E depends on the frequency of the incident light, and implies that photo-electric emission from a given metal will not take place unless the frequency of the light is high enough to cause $h\nu$ to exceed the value of p for that metal. This is found to be the case. For many metals the threshold value of ν, below which no emission can take place, is above the frequency of visible light. In such cases only ultra-violet has a sufficiently high frequency to produce any photo-electric emission. In some metals, notably cæsium, the threshold frequency is in the infra-red.

Further evidence of the particle nature of light is afforded by the Compton effect, in which the frequency of X-rays undergoes a change when they are scattered by electrons. The amount of the change can be worked out by regarding the X-ray beam as a stream of photons and applying the laws of mechanics to collisions between the photons and the electrons which they strike.

As a final example of the apparent particle nature of radiation we may mention the **scintillation counter.** In the early days of the study of radioactivity α-particles were detected and counted by the small flash of light which each made when it struck a specially prepared screen. Much more recently it has been discovered that γ-rays, whose nature is identical with that of light (although their wave-length is much shorter), produce

similar scintillations when they enter certain crystals. Each small flash represents the arrival of a γ-ray photon. The intensity of the γ-ray beam can be measured by arranging for the light from the scintillations to operate a sensitive photo-cell, which therefore delivers a short electric pulse to a counting circuit every time a photon arrives in the crystal.

It must not be supposed that we must necessarily abandon the conception of light as a wave motion because it behaves like a stream of particles in certain circumstances. It has become a fundamental part of present-day physical theory that the *two* characters are inherent in the nature of light and are separately revealed by different types of experiment.

Moreover, it is not only light and similar radiations which exhibit this dual character. Electrons, atoms, molecules and neutrons, which had hitherto been regarded simply as particles, have been shown experimentally to possess wave characteristics. The first demonstration of this occurred when it was shown that a beam of electrons which was reflected by a crystal or transmitted by a thin sheet of metal produced a diffraction pattern when it fell on a photographic plate. This behaviour is akin to that of a beam of X-rays under similar circumstances. When the wave-length associated with an electron was calculated from the dimensions of the pattern, it was found to be smaller than that of X-rays, and it is smaller still in the case of more massive particles such as neutrons and atoms. We have previously mentioned the electron microscope in which streams of electrons are used in an analogous way to light rays in an optical microscope. The resolving power of an electron microscope would be unlimited were it not for the wave character of the electrons, *i.e.* their property of forming diffraction patterns. As it is, the very small wave-length causes the resolving power to be 100 or more times greater than that of an optical microscope.

It has been established, therefore, that matter has a wave character, and a theory of **wave mechanics** has been developed accordingly. This theory accounts for small-scale effects, such as the diffraction of beams of particles, and at the same time it gives the laws of ordinary mechanics as approximations which apply to large-scale phenomena. Ordinary mechanics bears the same relation to wave mechanics as geometrical optics does to the wave theory of light, which, as we have seen, explains diffraction and interference and gives the principle of rectilinear propagation (on which geometrical optics is based) as an approximation suitable for large-scale effects. On account of the very short wave-lengths associated with material particles, their wave-effects are made evident only by very delicate experiments and, apart from these, it is usually adequate to describe their behaviour in terms of particle mechanics. The longer wave-length of light causes it to exhibit wave characteristics more readily than do material particles, and for long radio waves the wave behaviour is predominant.

EXAMPLES LV

1. If the wave-length of one of the sodium lines is $5{\cdot}89 \times 10^{-5}$ cm., and the speed of light is $3{\cdot}00 \times 10^{10}$ cm. sec.$^{-1}$, what is the frequency of the radiation? What changes, if any, would occur in the values of these quantities if the light, supposed originally to be travelling *in vacuo*, entered glass of refractive index 1·50?

Explain how matter may be excited by means of an electric current to give (*a*) a bright line spectrum, (*b*) a continuous spectrum. (L.Med.)

2. In experiments such as those for the determination of the velocity of light, very short intervals of time have to be measured. Explain two methods of measuring such short-time intervals. (L.H.S.)

3. Describe in detail a good terrestrial method of determining the velocity of light.

In an experiment by Foucault's rotating-mirror method the distance from the source to the rotating mirror is 1 m., that from the rotating mirror to the fixed mirror is 40 m. The speed of rotation of the mirror is 210 revolutions per sec. The image is displaced through 0·7 mm. Calculate the speed of light. (C.H.S.)

4. Describe some form of spectrometer. Describe, in the correct order, the adjustments which are necessary before the instrument can be used to measure the wave-length of a spectral line.

The emission spectrum of a certain element, suitably excited in the laboratory, consists of a number of sharp bright lines in definite positions. It is observed that (*a*) the sun's spectrum is crossed by dark lines occupying exactly these positions; (*b*) the bright lines can be identified in the spectrum of a distant star, but each shifted from its normal position towards the red end of the spectrum. Explain both of these observations as fully as you can. (O.H.S.)

5. Explain carefully how Huygen's construction enables the successive stages of the refraction of a plane wave-front at a plane boundary between two transparent media to be followed. Obtain an expression for the ratio of the refractive indices of the two media in terms of the velocities of light in the two media.

Give a concise account of the chief experimental evidence favouring the wave theory of light, as opposed to the corpuscular or emission theory. (O.H.S.)

ANSWERS TO EXAMPLES

Examples XLII. Page 708

1. (*a*) $180° - 2i$.

Examples XLIII. Page 737

1. Virtual object $8\frac{4}{7}$ cm. from mirror, upright.
2. 30 cm. from object, 120 cm., concave.
3. 12 cm.
4. (*a*) $u=15$ cm., $v=-7{\cdot}5$ cm.; (*b*) $u=-7{\cdot}5$ cm., $v=15$ cm.
5. 24 cm.
6. 0·11 in.
7. 12 cm., 3, concave.
8. 71·7 cm.
9. 4 mm. behind cornea, 1 mm.
10. 4 ft., 4.
11. 20 cm.
12. 15 cm.
13. (*a*) $8\frac{4}{7}$ cm. in front of mirror, (*b*) 60 cm. behind.
14. (*a*) 3 cm., (*b*) 1 cm.
15. $(2\sqrt{3}-3)$ in.

Examples XLIV. Page 771

1. 21·5 cm.
2. $180° + 2i - 4r$, 137·7°.
3. 75° 18′.
4. 1·64, 50°.
5. 9°.
6. 64°.
7. 56°.
8. 18° 36′.
9. 83° 38′.
10. 1·67.
13. 1·53.
14. 1·59, 75°.
15. 48′.
16. (*a*) 5°, (*b*) 5·6°.
17. 1° 1′.
18. 1·7°.

Examples XLV. Page 812

2. Diverging $3\frac{1}{3}$ in., 20.
3. $-3\frac{1}{3}$D.
4. (i) 200 cm. from lens, real, 8, (ii) 20·6 cm. from lens, virtual, 0·82.
5. 24 cm., 27 cm., 54 cm.
6. (*a*) 4 cm., (*b*) 18 cm. from lens on same side as object.
7. $\frac{5}{3}$.
8. $1\frac{7}{8}$ cm. in front of convex mirror, 10 cm. in front of mirror, 5 cm. beyond lens.
9. 3 in.
10. $\frac{f}{x}$, 10 cm. and 20 cm.
11. (*a*) convex 20 cm., concave 45 cm.; (*b*) $11\frac{1}{4}$ cm.
12. 2 cm. away from lens.
13. $27\frac{1}{4}$ cm. from lens.
14. $\frac{f}{7}$.

Examples XLVI. Page 833

1. 19·4 cm.
3. 26 cm. (diverging), 86·7 cm. (converging).

Examples XLVII. Page 843

1. Converging, 0·02D.
3. (*a*) 24 cm., (*b*) 9/4.
4. Converging, *R*.
5. Convex 15 cm., concave 22·5 cm., 10 cm.
6. Convex 12 cm., concave 21 cm., 6 mm., 14 mm.
7. 60 cm.
8. 12 cm.
9. (*a*) 156 cm., (*b*) 99·8 cm. and 64·0 cm.
10. 1·44.
11. 1·4.
12. $\frac{4}{3}$.
13. 1·6.
14. (*a*) $11\frac{1}{8}$ cm., (*b*) $5\frac{5}{8}$ cm., $3\frac{19}{27}$ cm.
15. 24 cm.
16. 1·43.

Examples XLVIII. Page 862

2. $-0·5$D, $22\frac{2}{9}$ cm.
3. Converging 2·75D, 36·4 cm.
4. 40–200 cm.
5. 100 cm., -1D, 25 cm.
6. 50 cm. parallel, 25 cm. perpendicular.
7. (i) $+4·2$D, (ii) $+\frac{2}{3}$D, 28·5 cm. in front of eye.

Examples XLIX. Page 895

1. 26, if image is at near point.
4. $\frac{2}{5}$, 30 cm. beyond convex lens.
6. $1\frac{7}{23}$ in. from objective, 23.
7. 1·05 cm. from objective.
8. 0·2 mm. away from object.
9. (*a*) 1·08 cm., 77.
10. $3\frac{3}{4}$ cm. from objective, $\frac{18}{9}$.
11. 27.
12. (*a*) image same size and position as object, (*b*) image at infinity, angular magnification 4·9, (*c*) blur, (*d*) inverted image at near point, angular magnification 5.
13. 7·7.
14. (*a*) 12, (*b*) 14·4.
15. 30 cm. from eyepiece, real, 2·33 cm. diameter.
16. (i) 5, (ii) 6 cm.
17. (*a*) 4, (*b*) 4·8.
18. 22·5 cm. from diverging lens, $\frac{31}{8}$.
19. $2\frac{4}{17}$ in., $\frac{34}{15}$.
20. 22·5 cm.

Examples L. Page 921

1. (*a*) 98 c.p., (*b*) $3\frac{1}{16}$ ft.-candles.
2. 650 c.p.
3. 2·52 ft.-candles, 0·28 ft.-candles.
4. 2·4 ft.-candles.
6. 5·05 ft.-candles, 2·57 ft.-candles.
7. 30 cm., 6 cm.
8. 0·5.
10. $\frac{105^3}{100^2}$ cm. from photometer.
11. 95 per cent.
12. (*a*) 4 ft.-candles, (*b*) 20 ft.-candles.
13. 333 m.-candles. (Solid angle in a cone of semi-angle θ is $2\pi(1 - \cos\theta)$.)
14. 2 ft.-candles, 10 ft.-candles.

Examples LI. Page 939

2. $1{\cdot}05 \times 10^{-5}$ cm.

Examples LII. Page 957

1. 1·08′.
2. 4062.

Examples LIII. Page 986

2. $1{\cdot}2 \times 10^3$ km. sec.$^{-1}$.

Examples LIV. Page 1000

2. 69°.

Examples LV. Page 1018

1. $5{\cdot}1 \times 10^{14}$ c.p.s., wave-length $3{\cdot}92 \times 10^{-5}$ cm., velocity $2{\cdot}00 \times 10^{10}$ cm. sec.$^{-1}$.
3. $3{\cdot}02 \times 10^{10}$ cm. sec.$^{-1}$.

INDEX TO VOLUME IV